
EÖTVÖS LORÁND TUDOMÁNYEGYETEM

TERMÉSZETTUDOMÁNYI KAR

BIOLÓGIAI INTÉZET

GENETIKAI TANSZÉK



Gímszarvasok egyedazonosítása és populációgenetikai vizsgálata autoszómás tetranukleotid mikroszatellita markerekkel

DOKTORI ÉRTEKEZÉS

Készítette:

Szabolcsi Zoltán

BIOLÓGIA DOKTORI ISKOLA

Iskolavezető: Dr. Erdei Anna, MTA rendes tagja

KLASSZIKUS ÉS MOLEKULÁRIS GENETIKA DOKTORI PROGRAM

Programvezető: Dr. Orosz László, MTA rendes tagja

Témavezető: Dr. Orosz László, MTA rendes tagja

Budapest

2013

Tartalomjegyzék

RÖVIDÍTÉSEK JEGYZÉKE.....	4
1 BEVEZETÉS.....	5
1.1 IRODALMI ÉS TÖRTÉNETI ÁTTEKINTÉS	5
1.1.1 A nem-humán igazságügyi genetika célkitűzései	5
1.1.2 A nem-humán igazságügyi genetika rövid története	5
1.1.3 Vadvilág igazságügyi genetika.....	6
1.1.4 STR markerek.....	7
1.1.5 STR allélok nevezéktana.....	9
1.1.6 A humán és nem-humán igazságügyi genetikai vizsgálatok statisztikai interpretációjának nemzetközileg elfogadott ajánlásai	11
1.1.7 A gímszarvas és annak hazai jelentősége.....	13
1.2 CÉLKITŰZÉSEK	16
2 ANYAGOK ÉS MÓDSZEREK	19
2.1 POPULÁCIÓS MINTAVÉTEL.....	19
2.2 MOLEKULÁRIS BIOLÓGIAI MÓDSZEREK	19
2.2.1 DNS izolálás	19
2.2.2 DNS koncentráció meghatározása	19
2.2.3 Gímszarvas mikroszatellita ortológok azonosítása a „zoo-kloning” eljárással	20
2.2.4 Gímszarvas mikroszatelliták polimorfizmus szelekciója	23
2.2.5 PCR fragmensek meghatározása, elválasztása és szelektálása kapilláris gélelektroforézissel	25
2.2.6 Multiplex rendszerek létrehozása.....	25
2.2.7 STR Allélok szekvencia sorrendjének meghatározása.....	27
2.2.8 Populációs egyedek genotipizálása a DeerPlex I-II rendszerekkel.....	27
2.2.9 A DeerPlex I-II rendszerek fajspecifikusságának tesztelése	28
2.3 POPULÁCIÓ GENETIKAI ÉS STATISZTIKAI MÓDSZEREK	29
2.3.1 Hardy-Weinberg egyensúly (HWE) teszt.....	29
2.3.2 Kapcsoltsági egyensúly (LE) teszt.....	30
2.3.3 Bonferroni korrekció.....	30
2.3.4 Populáció statisztikai fogalmak	31
2.3.5 A populáció genetikai struktúrája.....	34
3 EREDMÉNYEK ÉS KIÉRTÉKELÉSÜK	38
3.1 A „ZOO-CLONING” ALKALMAZÁSA A GÍMSZARVAS MIKROSZATELLITA MARKEREK AZONOSÍTÁSÁHOZ ÉS SZEKVENCIA ANALÍZISÉHEZ.....	38
3.2 GÍMSZARVAS TETRAUKLEOTID MIKROSZATELLITA LOKUSZOK SZELEKTÁLÁSA A POLIMORFIZMUS TESZT ALAPJÁN	40
3.3 A DEERPLEX I-II MULTIPLEX PCR RENDSZEREK	40
3.4 ALLÉLPOLIMORFIZMUSOK FELMÉRÉSE A DEERPLEX I-II RENDSZEREKKEL.....	44
3.5 A KIVÁLASZTOTT FRAGMENTEK SZEKVENCIA-ANALÍZISE.....	46
3.6 A POPULÁCIÓ GÍMSZARVAS EGYEDEINEK GENOTIPIZÁLÁSA DEERPLEX I-II RENDSZEREKKEL	46
3.7 A 10 MIKROSZATELLITA LOKUSZ JELLEMZÉSE.....	47
3.8 A DEERPLEX I-II MARKEREINEK FAJSPECIFIKUSSÁGA	48
3.9 HARDY-WEINBERG EGYENSÚLY TESZT	50
3.10 KAPCSOLTSÁGI (LE) TESZT ÉS MARKER LOKALIZÁCIÓ	51
3.11 A 10 AUTOSZÓMÁS STR LOKUSZ GENETIKAI DIVERZITÁSA	56
3.12 GENETIKAI DIFFERENCIÁLÓDÁS ÉS BELTENYÉSZET	59
3.13 A DEERPLEX I ÉS II INDIVIDUALIZÁLÓ EREJE	63
3.14 ESETTANULMÁNY.....	73
4 ÚJ MEGÁLLAPÍTÁSOK ÉS KÖVETKEZTETÉSEK	76
4.1 GÍMSZARVAS TETRAUKLEOTID MIKROSZATELLITA LOKUSZOK	76
4.2 POLIMORF GÍMSZARVAS TETRAUKLEOTID MIKROSZATELLITA MARKEREK.....	77

4.3	GENOTÍPIZÁLÁS <i>DEERPLEX I-II</i> MARKEREKKEL	78
4.4	A <i>DEERPLEX I-II</i> RENDSZEREK FAJSPECIFIKUSSÁGÁNAK VIZSGÁLATA	79
4.5	A <i>DEERPLEX I-II</i> MARKEREK FÜGGETLENSÉGÉNEK ÉS KROMOSZÓMÁS LOKALIZÁCIÓJÁNAK VIZSGÁLATA	80
4.6	A POPULÁCIÓK GENETIKAI DIVERZITÁSA ÉS GENETIKAI ÖSSZETÉTELE	82
4.7	A <i>DEERPLEX I-II</i> EGYEDEK ELKÜLÖNÍTÉSÉRE VALÓ ALKALMASSÁGÁNAK VIZSGÁLATA	83
5	KÖSZÖNETNYILVÁNÍTÁS.....	85
6	IRODALOMJEGYZÉK	86
7	FÜGGELÉK	97
8	ÖSSZEFOGLALÁS	120
9	SUMMARY	123

Rövidítések jegyzéke

AMOVA	molekuláris variancia analízis
α	szignifikancia szint
BLAST	Basic Local Alignment Search Tool
bp	bázispár
cDNS	mRNS-el komplementer DNS
DNY	délnyugat-magyarországi mintapopuláció
EK	északkelet-magyarországi mintapopuláció
ENO1	alfa enoláz 1
F_{IS}	populáción belüli beltenyészet paraméter
F_{IT}	átfogó beltenyészet paraméter
F_{ST}	Wahlund paraméter
h	Nei-féle géndiverzitás
H_0	megfigyelt heterozigóta gyakoriság
H_E	várt heterozigóta gyakoriság
HWE	Hardy-Weinberg egyensúly
in/del	inszerció/ delécio
ISFG	International Society for Forensic Genetics
LBM	valószínűség alapú módszer
LD	Linkage Disequilibrium – kapcsoltsági egyensúlytalanság
LE	Linkage Equilibrium – kapcsoltsági egyensúly
LR	valószínűségi hányados
nt	nukleotid
PAGE	poliakrilamid-gélelektroforézis
PCR	Polymerase Chain Reaction
PD	igazságügyi megkülönböztető erő
PIC	polimorfizmus információ tartalom
PI_{ave}	átlagos azonosság valószínűség
pM	egyezési valószínűség
p_{min}	minimális allélgyakoriság
SD	szórás
STR	Short Tandem Repeat
SWGDM	Scientific Working Group for DNA Analysis Methods
T_{an}	anneálási hőmérséklet
T_m	olvadási hőmérséklet

1 Bevezetés

1.1 Irodalmi és történeti áttekintés

1.1.1 *A nem-humán igazságügyi genetika célkitűzései*

Az igazságügyi genetikai vizsgálatok alapvető kérdése annak a megállapítása, hogy a törvényszéki eljárás során bizonyítékként szolgáló biológiai nyom milyen biztonsággal eredeteztethető egy adott fajtól, annak egy egyedétől vagy populációjától. Ez csak akkor lehetséges, ha az adott nyomból ki tudunk mutatni olyan tulajdonságot, amely jellemezhet egy adott fajt, annak egy populációját vagy egyedét. Ezeket a tulajdonságokat az élőlények örökítőanyagában kereshetjük. A DNS polimorfizmusnak köszönhetően, a genom minden pontjában megegyező két élőlény nem élhet a Földön, azonban a kérdésfeltevés vonatkozásában (ma még) nem vizsgálhatjuk meg a teljes örökítőanyagot. Az adott élőlény genomjának rendkívüli sokféleségéből (polimorfizmusából) egyfajta „mintát véve” annyi polimorf DNS markert vizsgálunk amennyi megfelelő erővel választ adhat a kérdéseinkre. A mi kérdésfeltevésünk az volt, hogy az általunk kiválasztott polimorf markerek individualizáló ereje elegendő-e a biológiai nyom adott egyedtől származtatásához.

1.1.2 *A nem-humán igazságügyi genetika rövid története*

Már a polimeráz láncreakció (PCR) felfedezésének évében, 1985-ben Alec Jeffreys és csoportjának köszönhetően került sor a DNS-vizsgálatok első törvényszéki alkalmazására. Dr. Jeffreys munkássága alatt rájött arra, hogy a genom nagyon sok igen variábilis tandem ismétlődést (ún. miniszatelliták, VNTR) tartalmaz, amelyek megfelelő polimorfizmust mutatnak ahhoz, hogy egy egyed ún. „genetikai ujjlenyomata” lehessen (Jeffreys *et al.*, 1985ab). A miniszatelliták vizsgálatához RFLP-alapú (restrikciós hossz-polimorfizmus) technológiát alkalmazott. Ebben az évben fedezték fel a polimeráz láncreakciót (PCR), amely kiküszöbölte a Jeffreys-próba legnagyobb hátrányát, ugyanis a DNS templát sokszorozásának köszönhetően jóval kevesebb genomi DNS is elegendő volt egy vizsgálathoz. 1987-ben a Jeffreys módszert madarakra is alkalmazhatóvá tették

(Burke *et al.*, 1987), majd Thommasen-nek köszönhetően 1989-től a nem-humán igazságügyi genetikában általánosan elterjedt egyedazonosító módszer lett (Thommasen *et al.*, 1989). A Jeffreys-féle eljárást fajmeghatározásra is alkalmazhatóvá tették, amelyet mitokondriális DNS markeren először Cronin végzett 1991-ben, majd Parson 2000-ben kidolgozta a fajazonosítás szekvencia meghatározáson alapuló módszerét (Cronin *et al.*, 1991; Parson *et al.*, 2000).

Törvényszéki eljárásokban először 1990-ben kezdték el markerként használni az emberi mikroszatellitákat (STR), 1994-ben pedig már megjelent az első humán STR-ket tartalmazó multiplex PCR-kit (Butler, 2011), amely előrelendítette a nem emberi STR markerek fejlesztését. Az első nem-humán multiplex rendszereket háziállatokra dolgozták ki. Stockmarks® (Applied Biosystems) kit néven az 1990-es évek végétől már elérhetőek voltak. Szarvasmarhára, kutyára és lovakra kidolgozott formában jelenleg is kereskedelmi forgalomban van. Hazánkban elsőként Pádár és munkatársai végeztek olyan nem-humán eredetű - kutya STR markereket alkalmazó - vizsgálatokat, amelynek eredményét bűnügyek megoldására is felhasználták (Pádár *et al.*, 2001, 2002). Kutyák egyedazonosítására ma már a nagyon szigorú nemzetközi előírásoknak (SWGDM) is megfelelő multiplex PCR kit is elérhető (Wictum *et al.*, 2013). Egy adott biológiai nyom fajösszetételének vizsgálatában a real-time PCR technológiát követik, Kanthaswamy és munkatársai 2012-ban kidolgozták ember-kutya-macska fajkeverékből álló minták mennyiségi analízisét is (Kanthaswamy *et al.*, 2012).

1.1.3 Vadvilág igazságügyi genetika

Jelen munka célkitűzései jól beilleszthetők a nem-humán igazságügyi genetikának egyik ágába, a vadon élő fajok védelmével foglalkozó vadvilág igazságügyi genetikába (ún. *wildlife forensic genetics*; Ogden *et al.*, 2009). A védett fajokkal történő illegális kereskedelem, az orvvadászat és más típusú természetkárosító tevékenységek nagy veszélyt jelentenek környezetünk hosszú távú megőrzésére. A vadvilág igazságügyi genetika a természetkárossal kapcsolatos bűnügyek megoldásán keresztül segíti a természeti kincseink védelmét. Különösen nagy hangsúlyt fektet a Washingtoni egyezmény által kategorizált veszélyeztetett fajok védelmére (CITES – Washington Convention on International Trade in Endangered Species of Wild Fauna and Flora). E

tudományágban a markerek fejlesztésének fő iránya a fajmeghatározás, az egyedazonosítás és az egyedek populációs eredetének meghatározása. A fogságban tartott és szaporított védett állatfajok esetében szükség lehet rokonsági vizsgálatokra is. Ez főleg akkor kérdéses, ha az adott védett faj tartása vagy kereskedelme kizárólag a vad populációkból való származás esetében tiltott. A *wildlife forensic genetics* molekuláris genetikai eszköztára alapvetően a konzerváció genetikai kutatásokban alkalmazott módszereken alapul, de a DNS bizonyítékok statisztikai interpretációja a humán igazságügyi genetikában használt alapelveket követi (Ogden et al., 2009). Vadon élő állat- és növényfajokra az utóbbi két évben egyre növekvő számban jelennek meg bűnügyekben is használható egyed-, vagy populációs származtatásra alkalmas több mikroszatellita markert tartalmazó (multiplex) PCR rendszerek (van Asch et al., 2010; Caratti et al., 2010; Lorenzini et al., 2011; White et al., 2012; Andreassen et al., 2012). Farkasokra és muflonokra specializált multiplex STR rendszerekkel a törvényszéki eljárások során több orvvadászati esetet is megoldottak már (Caniglia et al., 2010; Barbanera et al., 2012). A növekvő számú marker és módszer standardizálása érdekében a Nemzetközi Igazságügyi Genetikai Társaság (ISFG) 2012-ben több ajánlást tett közzé, amelyeket lehetőségeinkhez mérten mindig szem előtt tartottunk (Linacre et al., 2011).

1.1.4 STR markerek

A genom 150 bázispárnál kisebb magasan repetitív elemeket tartalmazó tömbjeit mikroszatellitáknak vagy rövid ismétlődő egységeknek (STR) nevezzük. Általában 2-6 bázispárból álló DNS szakaszok (ún. repeat egységek) ismétlődnek tandem módon egymásután. Az ismétlődő egységeket határoló szekvenciák (ún. flanking régiók) veszik körül, amelyek meghatározzák az adott tandem módon ismétlődő egység kromoszóma lokalizációját. Az STR lokuszok ismétlődéseinek száma az egyedek között nagy változatosságot mutat, és ez a polimorf tulajdonság teszi őket alkalmassá genetikai alapú egyedazonosításra (Butler, 2011). Az STR lokusz különböző alléljai általában méretükben különböznek egymástól, ezért az STR polimorfizmusokat a hossz-polimorfizmusok közé sorolják.

A humán genomban polimorf mikroszatelliták ezreit azonosították. Az emberi 23-as

kromoszóma több mint 8000 STR-t tartalmaz. Az állat-, és növényfajok esetében a mikroszatelliták izolálásának legáltalánosabb módszere az ismert genom szekvenciájú rokon fajokból nyert mikroszatelliták adoptációja (*Dawnay et al., 2008*) vagy az ismeretlen genomból biotechnológiai úton nyert izolálás (*Jones et al., 2000*). A mai újgenerációs szekvenálási módszerekkel indított genom programok során egyre több fajból származó STR markert fedeznek fel.

Az STR ismétlődő szekvenciákat repeat egységük hosszának megfelelően nevezik el: a di-, tri-, tetra-, penta-, hexa-repeat egységek 2, 3, 4, 5, 6 nukleotidból állnak. Napjainkban a humán igazságügyi genetikában elsősorban a tri-, tetra és pentamer repeatekből álló mikroszatellitákat használják, a nem-humán vizsgálatokban azonban még elsősorban dimer repeatekből álló markerek (dimer mikroszatelliták) a dominánsak.

Az STR szekvenciák nem csak repeat-ek számában, hanem struktúrájában is variabilitást mutatnak. Az egyszerű repeat régióval rendelkező mikroszatelliták azonos hosszúságú és szekvenciájú egységekből állnak. Az összetett struktúrájuk két vagy több különböző egyszerű, egymással szomszédos repeat-et tartalmaznak. A komplex STR-ek esetében a különböző szekvencia összetételű és változó hosszúságú repeat-egységeket változó méretű közbeiktatott szekvencia szakítja meg. A komplex hipervariábilis STR-ek ezen felül igen nagy változékonyságot mutatnak.

Nem minden STR allél mérete felel meg az ismétlődő egységek egész számú többszörösének. Vannak olyan, ún. intermediér (köztes, mikrovariáns) allélok is, amelyek a repeat-ben vagy a flanking régióban egy vagy több nukleotidot ért megváltozást hordoznak, s emiatt ezek mérete a konszenzus allél méretével nem egyezik meg (*Butler, 2011*). Előfordulhatnak azonban olyan mikrovariáns allélok is, amelyek méretükben megegyeznek, azonban az azonos számú, de eltérő DNS szekvenciájú repeatek miatt vagy a flanking régiók szekvencia eltérése következtében különböznek (*Wang et al., 2012*).

Az STR markerek igazságügyi célú bevezetését nagyban elősegítette, hogy a határoló szekvenciájukra tervezett és hibridizálódó oligo (primer) szekvenciák segítségével a PCR reakció során felsokszorozhatók. Ennek következtében a biológiai anyagmaradványokból igen sok információ nyerhető, legyen szó annak egyedi származásáról vagy faji összetételéről. Másik igen kedvező tulajdonsága, hogy egy

reakció során egyidejűleg több DNS-lokuszt is felsokszorozható (ez az ún. multiplex PCR), valamint a felsokszorozandó DNS szakaszok (amplikonok) lerövidítésével igen leromlott minőségű és töredezett (degradált) DNS tartalmú minták is vizsgálhatóvá válnak. A multiplex PCR e tulajdonságainak köszönhetően akár néhány sejt genetikai állománya is elegendő lehet az egyed-, vagy fajazonosítás számára. A munkánk során azt szerettük volna elérni, hogy legalább 5 nanogrammmal gímszarvas eredetű örökítő anyag (azaz kb. 800-1000 sejt DNS-tartalma) elegendő legyen az egyedek azonosítására kidolgozott rendszerünkhöz.

A jelenlegi elfogadott gyakorlat szerint az STR allélok DNS-fragmens analízissel az esetek döntő többségében egyértelműen elkülöníthetők egymástól. Az STR lokuszok multiplex PCR-alapú szimultán vizsgálata lehetőséget teremtett humán igazságügyi genetikában a személyazonosítás gyors kivitelezésére (*Kimpton et al., 1994*). Az 1990-es évek folyamán a mikroszatellitákat választották a személyazonosítás fő genetikai markeréül (*Butler, 2011*). A nem-humán és a humán igazságügyi gyakorlatban előírt követelmény a PCR reakció során alkalmazott primer szekvenciák keresztreagáló képességének vizsgálata, ugyanis ennek eshetősége megnehezítheti vagy bizonytalanná teheti az adott célorganizmus vizsgálatát. Egyedazonosítás céljából csak az adott faj STR lokuszának határoló régiójára specifikusan hibridizáló primer oligo szekvenciája használható (*Linacre et al., 2011*).

1.1.5 STR allélok nevezéktana

A nem-humán igazságügyi genetikai vizsgálatokban alkalmazott STR lokuszok és alléljaik nevezéktana az ISFG humán mikroszatellitákra kidolgozott ajánlásain alapul (*Gill et al., 1994, 1997a, 2001; Bär et al., 1997; Olaisen et al., 1998*). A komplex struktúrájú STR lokuszok alléljainak elnevezésében az Eichmann és munkatársai által kidolgozott módszer javasolt (*Eichmann et al., 2004*).

A legfontosabb nem-humán igazságügyben alkalmazható szabályok:

- A szekvenciák leolvasása mindig 5'-3' irányban történhet, és a tipizáláshoz használatos DNS szál az kell, hogy legyen, amelyet a nyilvános adatbázisban először publikáltak.

Az allélikus nevezéktanra vonatkozó ajánlások fő pontjai:

- Az elnevezés a teljes ismétlődési egységek (repeatek) száma szerint történik, amely egyszerű összetételű STR markereknél könnyen alkalmazható.
- Amennyiben az ismétlődő régió nem teljes egységet (repeat) is tartalmaz (mikrovariánsok, köztes (intermedier) allélok), úgy a teljes repeatek száma után egy ponttal elválasztva kell feltüntetni a hiányos ismétlődő egység meglévő bázisainak számát.
- Az újonnan felfedezett STR-ek alléljait a variáns és nem-variáns ismétlődési egységek együttes számának alapján kell elnevezni.
- Komplex STR-ek esetében, amelyek eltérő szekvenciájú és többféle méretű repeatekból állnak, több alternatív nevezéktan is létezik. Ennek alapját az ismétlődő régió hosszának definiálása jelenti. A régió nukleotidokban kifejezett hosszát ezután osztani kell a tetramer ismétlődő egységek számával. Ez azonban túl sok mesterséges mikrovariáns allélt eredményezhet, ezért az Eichmann-féle módszer szerint a definiált ismétlődő régiót úgy kell módosítani, hogy az minél kevesebb intermedier elnevezést eredményezzen (*Eichmann et al., 2004*).
- Köztes allélok az ún. flanking (határoló) régiók egyszerű inszerciós/deléciós mutációjának köszönhetően is megjelenhetnek (*Füredi et al., 1999*), ahol az eredeti primereket használva N.1 és N.2 köztes allélok keletkeznek, amelyeket a fent említett esetektől való megkülönböztetés végett + illetve - jellel ajánlatos jelölni. A + jel inszercióra, a - pedig delécióra utal.
- Köztes allélok keletkezhetnek flanking régió nukleotidjainak egyszerű transzverziós vagy tranzíciós mutációjával is. Ebben az esetben a mutációs pont ismétlődő egységtől számított helye jelölendő. Például a humán Y-kromoszómás DYS391 lokusz repeat régiójától downstream 87 nukleotid távolságra detektált C-G mutációjának jelölése (D87C→G).

Az emberi és nem-emberi igazságügyi genetikában alkalmazott mikroszatelliták standardizált elnevezésű alléljai, DNS szekvenciái és a lokuszok kromoszómás elhelyezkedése bizonyos fajok esetében a nemzetközi STRBase adatbázisban hozzáférhető (*Butler and Reeder, 1997*).

1.1.6 A humán és nem-humán igazságügyi genetikai vizsgálatok statisztikai interpretációjának nemzetközileg elfogadott ajánlásai

Az igazságügyi genetikában a DNS-bizonyíték statisztikai interpretációjának legáltalánosabban elfogadott módszere az ún. Bayes-elvű hipotézistesztelés (Ewett and Weir, 1998). A valószínűség számítás Bayes-tételének esély formájú leírása szerint a nyomozó hatóság, ill. a bíróság által az igazságügyi szakértőnek feltett alapkérdést: pl. *Mekkora a valószínűsége annak, hogy a helyszínen talált vérnyom egy adott egyedtől származott?* mindig legalább egy másik hipotézis valószínűségéhez kell viszonyítani: pl. *Mekkora a valószínűsége annak, hogy a vérnyom az adott populációból véletlenszerűen kiválasztott másik egyedtől származott?* Az első kérdés általában a vád feltevésére (H_p), a második kérdés pedig általában a védelem hipotézisére (H_d) vonatkozik.

Az előzőekben említett ún. *posterior* valószínűség további két valószínűségi értékre bontható. Az ún. *likelihood* tag a genetikai vizsgálatra vonatkozik. A *prior* valószínűség a DNS bizonyíték ismerete nélkül, az egyéb körülmények ismeretében (I) az adott kérdésre vonatkozó *posterior* valószínűségi értéket adja meg.

$$\frac{P(H_p|E, I)}{P(H_d|E, I)} = \frac{P(E|H_p, I)}{P(E|H_d, I)} \times \frac{P(H_p|I)}{P(H_d|I)}$$

A *likelihood* azt mutatja meg, hogy mekkora a DNS-vizsgálat eredményének valószínűsége a vád, ill. a védelem feltevésének (H_p és H_d) teljesülése esetén. A rendelkezésre álló DNS bizonyítékból (E) - pl. a gyanúsítottól lefoglalt agancstrófeából - kimutatott gímszarvas DNS profil (G_s) és a helyszíni nyomból kimutatott gímszarvas genetikai profil (G_c) - ismeretében a két *likelihood* hányadosa (LR):

$$LR = \frac{P(G_s, G_c|H_p, I)}{P(G_s, G_c|H_d, I)} = \frac{P(G_s|G_c, H_p, I)}{P(G_s|G_c, H_d, I)} \times \frac{P(G_c|H_p, I)}{P(G_c|H_d, I)},$$

ahol az egyenlet jobb oldalának második tagja a helyszíni profil populációs előfordulását kifejező valószínűség. A vád hipotézisének *likelihood* értéke: *Mekkora a valószínűsége annak, hogy éppen a vérnyomból kimutatott DNS-profilt határozzuk meg akkor, ha feltesszük, hogy a vérnyom az adott egyedtől származott?*, a védelem ellenhipotézisének valószínűsége, *Mekkora a valószínűsége annak, hogy éppen a vérnyomból kimutatott DNS-profilt határozzuk meg akkor, ha feltesszük, hogy a*

vérnyom az adott populációból véletlenszerűen kiválasztott másik egyedről származott? Fontos azt hangsúlyozni, hogy az ellenhipotézis *likelihood* értékének számítása csak akkor korrekt, ha a statisztikai kalkulációhoz felhasznált populációs allél-, vagy profilgyakorisági adatbázis típusát a nyomozati információ (pl. a bűncselekmény elkövetésének helye) határozza meg.

Amennyiben $G_s = G_c$, vagyis a genetikai profilok megegyeznek, a biológiai nyom csak az adott (al)faj egy egyedének genetikai anyagát tartalmazza valamint a laboratóriumi genotipizálás (profil kimutatás) hibája kizárható, akkor a valószínűségi hányados (LR) – legegyszerűbb formájában – a DNS-profil egyezési valószínűségének (pM) reciprokával egyezik meg.

$$LR = \frac{1}{P(G_s | G_c, H_d, I)} \approx \frac{1}{pM}$$

A pM azt fejezi ki, hogy mekkora annak az esélye, hogy a referencia egyeden kívül még egy olyan egyedet találunk a populációban, amelynek DNS-profilja teljes egyezést mutat a biológiai nyom genotípusával. Az 1990-es évek elejéig az egyezési valószínűség számításakor a szakértők azzal a feltételezéssel éltek, hogy a profilok függetlenek egymástól, vagyis az egyezést mutató profil megfigyelésének várt valószínűsége nem függ az első profil megfigyelésének valószínűségétől. Az egyezési valószínűség kalkulációja a Fisher-Wright modell alapján a Hardy-Weinberg (HW) törvényszerűségeknek megfelelően történt. Nem minden adatbázisban teljesülnek azonban a HW szabályok, ugyanis bizonyos populációs tényezők befolyásolhatják a populáción belüli allélok és genotípusok random párosodásából adódó egyensúlyát. A struktúráltság, vagyis az alpopulációk esetleges jelenléte a populáción belül a Wahlund hatás következtében a populáció heterozigóta genotípusainak túlbecslését és a homozigóta genotípusainak alulbecslését okozza. Az alpopulációkban élő egyedek között, amennyiben teljesülnek a HW kritériumok, a párválasztás valószínűsége nagyobb, mintha az, az alpopulációban elő és az azon kívüli egyedek között történne. Az alpopulációt alkotó egyedek között tehát bizonyos mértékű genetikai korreláció létezhet. Ennek következtében az adott alpopulációból származó egyed kalkulált profilgyakorisága nagyobb annál, mintha azt a teljes adatbázis allélgyakorisági értékeiből számítanánk. Balding és Nichols (*Balding and Nichols, 1994*) megoldást keresett az adatbázison belüli lehetséges alpopulációk meglétének problémájára. A

kutatók a profil egyezési valószínűséget megadó képleteikbe a közös leszármazási együtthatót (θ -theta faktor) építették be, hogy kiegyenlítsék a struktúráltság következtében kialakult Wahlund hatást. Amennyiben az alpopulációban a HW egyensúly feltétele teljesül akkor a $\theta = F_{ST}$. 1996-ban az USA Nemzeti Kutatótanácsa is javasolta (*National Research Council, 1996*) a számítási módszert, amelynek hatására ez a megközelítés a szakértői társadalomban elfogadottá vált (*Ewett and Weir, 1998*). Ayres és Overall (*Ayres és Overall, 1999*) rávilágított arra a problémára, hogy Balding és Nichols formulái azon a feltételezésen alapultak, hogy az alpopuláción belül a párválasztás véletlenszerű. Ez a feltétel azonban állati populációkban ritkán teljesül. A beltenyészet hatása hasonló a Wahlund effektushoz. A populáción belül a két hatás együttes fellépése esetén a $\theta = F_{IT}$. Ennek alapján Ayres és Overall tanulmányukban a profil egyezési valószínűség számítására egy olyan formulát javasolt, amely az F_{ST} mellett az F_{IS} (populáción belüli ún. beltenyésztési együttható) értékét is tartalmazta. A nem-humán igazságügyi genetikai bizonyítékok statisztikai interpretációjánál ez a formula a nemzetközi szinten elfogadott (*Linacre et al., 2011*).

1.1.7 A gímszarvas és annak hazai jelentősége

Rendszertani helyét tekintve a gímszarvas (*Cervus elaphus*) a *Cetartiodactylia* renden belül a Párosujjú patások (*Artiodactylia*) monofiletikus csoportjába, azon belül a Szarvasfélék családjába (*Cervidae*) tartozik. A *Cervidae* családot anatómiai bélyegek és molekuláris filogenetikai elemzések alapján 23 nemzetségre és 47 fajra osztható.

A jelenlegi taxonómiai álláspont szerint a család három alcsaládja (Újvilági szarvasfélék - *Capreolinae*, Óvilági szarvasfélék - *Cervinae*, Víziőzformák - *Hydropotinae*) közül a gímszarvas az Óvilági szarvasok (*Cervinae*) csoportjába sorolható. Az Óvilági (*Cervinae*) és Újvilági (*Capreolinae*) szarvasfélék az elülső lábak két szélső elcsökevényesedett lábközépcsontjainak helyzete alapján jól elkülöníthető. A *Plesiometacarpalia* csoportnál, ahová az Óvilági formák, így a hazai däm-, és gímszarvas is tartozik, a két szélső csökevényes lábközépcsont a lábtőcsontok ízületéhez kapcsolódik, míg az Újvilági formáknál ezek az ujjpercekhez ízesülnek. A szarvasfélék első tagját a Hindukus-hegységből származó 25 millió éves leletekből ismerjük. Ennek a két családnak az elválása már a késői Miocénben (7-8 millió éve)

elkezdődött, a szarvasfélék adaptív radiációjának kezdetén, amelyet a globális lehűlés és tengerszint csökkenés következtében kialakult nagy kiterjedésű füves területek okoztak. Ekkor érhatték el az Észak-Amerikai földrészt is. A Floridában talált *Eocoileus gentryorum* 5 millió éves leleteiből kiderült, hogy ekkor már éltek cervidák az amerikai földrészen. Később, a Pleisztocénben létrejött Panamai földhídon keresztül terjedtek el Dél-Amerikában. A *Cervus* nemzetség kb. 5 és 2 millió éve a Pleisztocénben jött létre. Az Óvilági csoporthoz tartozó vapitik ősei a Pleisztocénben kezdődött glaciális-interglaciális periódusok idején kelhettek át Ázsiából Beringián keresztül Észak-Amerikába. Erre az időszakra tehető az összes mai modern *Cervidae* fajok őseinek létrejötte is (Pitra et al., 2004; Groves and Grubb, 2011; Myers et al., 2013).

A *Cervus elaphus* faj Holarktikus elterjedésű, korábban 22 alfaját különböztették meg. A mai rendszerezők több faj/alfaj státuszát vagy meglétét módosították. A citokróm b gén mutációinak filogenetikai elemzése során kiderült, hogy a genetikai távolság *Cervus* nemzetség - meglepő módon a Thorold szarvast (*Przewalskium albirostris*) és a szikaszarvas (*Cervus nippon*) alfajait is magában foglaló - két kládra osztható (I-II.). A Tarimi csoport két tagja, a Jarkand-, és az Afgán gímszarvas (*Cervus e. yarkandensis*, *Cervus e. bactrianus*) a Nyugati klád csoportjához (I) sorolható, amelybe az Észak-afrikai, Európai és a Balkáni gímszarvasok tartoznak. A Balkáni csoport tagjai (*Cervus e. hippelaphus*) jól elkülönülnek a Nyugat-, és Észak-európai gímszarvasoktól. A *Cervus e. montanus* alfaj megléte erősen megkérdőjelezhető. A szikaszarvas alfajai a Thorold szarvassal együtt a Keleti kládba (II) tagozódnak. Ennek tagjait ma már nem alfajként, hanem külön fajként kezelik, így az Észak-Amerikai és Kelet-Ázsiai vapitiket is. Az európai gímszarvasok és a szika-, Thorold szarvasok szekvencia hasonlósága tehát igen nagy lehet (Mahmut et al., 2002; Polziehn and Strobeck, 2002; Pitra et al., 2004; Ludt et al., 2004).

A gímszarvas, mint nagytestű herbivor hazai faunánk ökológiai és vadgazdálkodási szempontból is nagy jelentőséggel bíró tagja. Vadászható nagyvad, amelynek létszáma emelkedő tendenciát mutat. Hazai állománya a 2012/2013-as vadászati idény teríték adatai alapján 100.000 egyednél többre becsülhető (Csányi et al., 2010). A trófea minősége szempontjából a magyarországi gímszarvas kiemelkedő más európai állományokhoz képest. A jelenlegi szabad területről származó trófea világranglista első

10 tagjának fele Magyarországról származik, köztük a méltán világhírű, 1986-ban elejtett Karapancsai gímszarva bika agancstrófeája, amely a lista harmadik helyezettje. Ezek a magyar rekorder trófeák a déli-, délnyugati országrészről származnak. A szarvas agancs kvalitatív morfológiai karaktereit vizsgálva Köller és munkatársai 1988-ban (Köller *et al.*, 1988) szignifikáns regionális különbségeket mutattak ki magyarországi gímszarvas populációk között, azonban ezek a vizsgálatok cáfolták a Szederjei-féle rasszok meglétét (Szederjei, 1965). Már korán felvetődött annak a lehetősége, hogy a gímszarvas agancs kvantitatív és kvalitatív jellegeinek állományok közötti különbségének genetikai háttere is lehet. Hartl és munkatársai már 1990-ben vizsgálták ennek lehetőségét (Hartl *et al.*, 1990). Polimorf izoenzimeknél (pl. izocitrát-dehidrogenáz, IDH) kismértékű, de szignifikáns alléleloszlás különbségeket mutattak ki magyarországi állományok (Zala megye, Somogy megye, Gemenc, Pilis, Börzsöny) között, a hazai démek lehetőségét tehát már ők is felvetették.

Az elhullásból és a legális elejtésből származó éves egyedszám csökkenés 2011/2012 vadászati szezonban hozzávetőlegesen 49.000 példány volt (Csányi *et al.*, 2010). Vadon élő gímszarvas állományunkra nagy veszélyt jelent az orvvadászat. Sajnos az illegális elejtésből származó veszteséget azonban csak becsülni lehet. Sugár (Sugár, 2002) szerint az országos viszonylatban az illegális elejtésből adódó veszteség 8-10 ezer egyedre tehető. Pechtol (Pechtol, 2001) szerint a felderített esetek száma 1999-2000-ben 252 volt, az okozott kár pedig több százmillió Forint. A hivatásos vadászok véleménye szerint minden harmadik nagyvad az orvvadászok zsákmánya lesz (Nagy, 2002). A vadászok által elloptott vad száma elérheti, vagy meg is haladja a hivatalosan közzétett éves teríték akár 50%-át is (Homonnay, 2004).

Az orvvadászat és vadlopás elleni védekezés egyre nagyobb probléma, a megoldás szüksége a jogalkotás szintjére is begyűrűzött. A Büntető Törvénykönyv 2012-es módosításának köszönhetően, amely 2013. július. 1-én lépett hatályba, az orvvadászat büntetőjogi kategória lett (Hatályos Jogszabályok Gyűjteménye, <http://net.jogtar.hu/>):

„245. § Aki

a) vadászterületen vadászatra való jogosultság nélkül, illetve idegen vadászterületen vadászsként engedély nélkül vad elejtésére vagy elfogására irányuló tevékenységet végez,

b) külön jogszabályban meghatározott, a vadfaj valamennyi egyedére kiterjedő

vadászati tilalmi idő hatálya alá eső vadfaj egyedét ejti vagy fogja el,

c) külön jogszabályban meghatározott tiltott vadászati eszközzel, tiltott vadászati módon vagy kíméleti területen vad, illetve fokozottan védett vagy védett gerinces állat elejtésére vagy elfogására irányuló tevékenységet végez,

bűntett miatt három évig terjedő szabadságvesztéssel büntetendő.”

A törvény adta lehetőségeket kiegészítve a munkánk során létrehozni kívánt genetikai alapú egyedazonosító rendszer megkönnyítheti a terepen dolgozó hivatásos vadászok munkáját is, ugyanis nem teszi szükségessé az orvvadászok sokszor igen nagy veszéllyel járó tettenérését. Az általunk megoldott orvvadászati eset jól példázza, hogy a bűncselekmény megalapozott gyanújához a tett helyszínéről biztosított minta is elegendő lehet.

1.2 Célkitűzések

Doktori munkámnak egyik célja az volt, hogy az orvvadászat elleni küzdelemben használható olyan eszközt hozzunk létre, amellyel az elkövetőt a tettenérés szüksége nélkül lehessen összefüggésbe hozni közvetlenül a tett helyszínével. A hatóság által biztosított vér-, vagy egyéb szövetmintákat a belőlük kimutatott genetikai profil alapján származtatni lehet a feltételezhetően illegális elejtésből származó gímszarvas trófeájából vagy húsból. A mikroszatellitákból álló genetikai profil individualizáló erejének felméréséhez szükséges populációs vizsgálatok úttörő jelleggel szolgáltak a hazai gímszarvas állomány populációgenetikai vizsgálatában is, ugyanis más szerzők által (*Hartl et al., 1990*) korábban nem vizsgált élőhelyekről származó egyedeket is bevonhattunk a populációs felmérésbe. A másik cél a gímszarvas magyarországi populációgenetikai felmérésének folytatása volt, ugyanis a populáción belül izolált szaporodási közösségekről (démekről) és azok térbeli lokalizációjáról nagyon keveset tudunk. Ezek meglétének kimutatása nemcsak a korábban említett igazságügyi genetikai vonatkozások miatt fontos. Vadbiológiai és vadgazdálkodási jelentőséggel is bír. A démek térbeli elhelyezkedéséhez igazított vadgazdálkodási egységek határai az ökológiai szemléletű vadgazdálkodás fontos részét képezik. Ennek megvalósulása esetén ugyanis a valós populációdinamikai paraméterek mérésével pontosabb hasznosítási ráták becsülhetők (*Csányi and Lehoczki, 2010*).

A gímszarvasok populációgenetikai vizsgálatában számos kiváló publikáció is megjelent már, amelyek az állományok genetikai struktúráját vizsgálták, a térben elkülönült populációk között géndiverzitás és alléleloszlás különbségeket mutattak ki valamint az izolált populációk között becsülték a génáramlás mértékét (*Kuehn et al., 2003, 2004*), ezekhez vizsgálatokhoz a szarvasmarha, juh és rénszarvas ismert dimer STR markereit adoptálták. Dinukleotid mikroszatellitákat alkalmaz a Zsolnai és munkatársai által rokonsági vizsgálatok céljából kifejlesztett 8 markert tartalmazó STR készlet is (*Zsolnai et al., 2009*). Bizonyos populáció genetikai kutatások már vegyesen alkalmazzák a di-, és tetramer markereket (*Pérez-Espona et al., 2008*). Orvvadászati esetek egyedek identifikálásával történő megoldásához először Jobin munkacsoportja alkotta az első szarvas egyedazonosításra alkalmas dimer mikroszatellitákból álló rendszerét 2008-ban (*Jobin et al., 2008*), amit 2010-ben a Sitka szarvasokra kidolgozott rendszer követett (*Brinkman et al., 2010*). A dimer markereket ismert genotipizálási nehézségeik következtében humán forenzikus célokra már nem használják (*Butler, 2011*). A multiplex PCR rendszer fejlesztésénél ezért a szakirodalomban rendelkezésre álló tetranukleotid egységeket tartalmazó szarvas markerekből indultunk ki (*Jones et al., 2000, 2002; Meredith et al., 2004*).

A tetranukleotid mikroszatellitákból álló, egyedazonosítás és populáció genetikai célokat szolgáló multiplex PCR rendszer létrehozására a következő munkafázisokat tűztük ki:

I. Az STR markerek igazságügyi hitelesítésének részét képezik a populációs felmérések. Az allélpolimorfizmusok meghatározásához populációs mintavételezésbe kezdtünk. Több olyan gímszarvas élőhelyet is be szeretnénk vonni a felmérésbe, amelyeket korábban más szerzők még nem vizsgáltak. A mintagyűjtést kiszélesítését terveztük az eddig még kevésbé vizsgált észak-magyarországi állományokra.

II. A specifikus PCR primerek tervezhetősége érdekében a szakirodalomban meglévő, öszvérszarvas (*Odocoileus hemionus*) és amerikai vapiti (*Cervus canadensis*) eredetű tetranukleotid mikroszatellita lokuszok gímszarvas ortológjainak meg kellett határozzuk a DNS szekvenciasorrendjét. Ehhez a csoportunk által kifejlesztett un. „zoo-

kloning” eljárást szándékoztuk alkalmazni.

III. A rendelkezésre álló gímszarvas markerek közül azokat kellett összeválogassuk, amelyek polimorfizmus foka elegendő az egyedek identifikálásához. A markerek allélszámát és azok gyakoriságát magyarországi populációkon még nem vizsgálták, ezért ezt fel kellett mérjük. Ennek első lépéseként a PhastSystem technológia alkalmazásával polimorfizmus alapú szelekciót terveztünk.

IV. A polimorf markerekből multiplex PCR rendszerek létrehozását terveztük.

V. A genotipizálást a humán igazságügyi genetikai vizsgálatokban is alkalmazott módszerekre alapoztuk. A markerek fluoreszcens festékkel történő jelölésével automatizált kapilláris-gélelektroforézis rendszert kívántunk bevezetni. A markerek jelöléséhez az öt színes DS-33 fluorofórijait kívántuk felhasználni, az allélok elválasztásához az ABI3130 kapilláris-gélelektroforézis készülék állt rendelkezésünkre, az allélméret meghatározásához pedig a GeneScan 500 LIZ standardot alapul véve a GeneMapper v3.2 szoftvert választottuk.

VI. A multiplex mikroszatellita készlet igazságügyi alkalmazhatósága érdekében meg kellett vizsgáljuk azt, hogy a kiválasztott polimorf lokuszok együttes identifikáló ereje elegendő statisztikai erővel rendelkezik-e ahhoz, hogy egy biológiai nyomot egy bizonyos egyedtől származtassunk. Ehhez szükséges a lokuszok Hardy-Weinberg egyensúlyának és azok független öröklődésének vizsgálata. El kellett végezzük a populáció genetikai összetételének vizsgálatát is, hiszen ennek befolyása van a több lokuszos genetikai profil identifikáló erejére.

VII. Orvvadászati esetekben nagy jelentőséggel bírhat, ha egy biológiai nyomnak nemcsak egyedi eredetét tudjuk valószínűsíteni, hanem az azt hátrahagyó egyed forráspopulációját is. Ehhez azonban meg kellett vizsgáljuk a populációs minták (alpopulációk) genetikai homogenitását ahhoz, hogy azok alkalmasak-e az egyedek populációs eredetének valószínűsítésére.

2 Anyagok és módszerek

2.1 Populációs mintavétel

A mintákat a korábbiakban más szerzők által is vizsgált (*Hartl et al., 1990*) északkelet-, és délnyugat-magyarországi gímszarvas állományokból (EK és DNY) gyűjtöttük kiegészítve őket a gödöllői-dombsági, cserhádi-, bükki- és mátrai állományokból származó egyedekkel (*Szabolcsi et al., 2008*). A minták származási helyét és típusát táblázatos formában közöltük (F.1., F.2. ábrák). Az EK és DNY állományokból összesen 100 db mintát vizsgáltunk, 48 db-ot az északkeleti és 52 db-ot a délnyugati állományból. A minták szöveti eredete vegyes volt, 56 db vérmintát, 37 db izommintát és 7 db hullott agancsot tartalmazott. A vérminták egy 5x5 cm méretű itatós papírlapon beszáradt állapotban érkeztek, amelyeket hivatásos vadászok gyűjtöttek az elejtett szarvasokból.

2.2 Molekuláris biológiai módszerek

2.2.1 DNS izolálás

Az itatós papírokon lévő vérfoltok kivágott darabjaiból a DNeasy Blood & Tissue Kit (QIAGEN®) felhasználásával DNS-t tisztítottunk. Az izom mintákból kb. 0,2 cm³-nyi darabot 500 µl SEB (DNS extrakciós puffer) 10 µl proteináz K (Fermentas, 20 mg/ml) enzimet tartalmazó oldatban egy éjszakát emésztettük (*Lincoln and Thomson, 1998*). Az oldatot ezután fenol/kloroform/izo-amilalkohol 25:24:1 arányú elegyével extraháltuk, majd Microcon YM-100 makromolekula szűrővel koncentráltuk. A szűrletet TE oldatban -20°C-on tároltuk.

A hullott agancsokból – a mintavételi felület fizikai tisztítása (csiszolás), etilalkoholos lemosása és UV fénnel történő sterilizálása után – fúróhegy segítségével kb. 200-250 mg csontport nyertünk ki. A csontporból a DNS kinyerését Kalmár és munkatársai által publikált módszerrel végeztük (*Kalmár et al., 2000*).

2.2.2 DNS koncentráció meghatározása

A minták teljes nukleinsav tartalmát Nanodrop 2000 UV-Vis spektrofotométer (Thermo Scientific) felhasználásával mértük meg. A készülék szoftvere a minták DNS koncentrációját az oldatok fényelnyelésére vonatkozó Lambert-Beer törvény alapján kalkulálta. Az összefüggés szerint $c = A \times \epsilon / b$; ahol c az oldat koncentrációja (M), A a mért abszorbancia egység (AU), ϵ az extinkciós koefficiens ($M^{-1}cm^{-1}$), és b a rétegvastagság (cm). A készülék egyik különlegessége az, hogy mérőedényt nem igényel, az oldatból 2 μ l-es cseppet kell egy speciális mintatartóba helyezni. A minták tisztaságát a mért abszorbancia hányadosokkal (A260/280, A260/230) ellenőriztük. Az olyan mintákból, amelyek A260/280 értéke 1,8 - 2,0 intervallum közé esett, valamint amelyek A260/230 értéke $>2,5$ ng/ μ l-es hígításokat készítettünk (*NanoDrop 2000/2000c Spectrophotometer V1.0., User Manual*).

2.2.3 Gímszarvas mikroszatellita ortológok azonosítása a „zoo-kloning” eljárással

A vadvilág igazságügyi genetikában a vadonélő állatfajokra alkalmazható genetikai markerek száma általában igen alacsony, ami a legtöbb faj teljes genomsekvenciájának hiányából fakad. Egy adott faj STR markereinek forrása legtöbbször a más rokon fajokból történő adoptáció vagy a biotechnológiai módszerekkel végzett STR izoláció (Jones *et al.*, 2000; Dawney *et al.*, 2008). A mi esetünkben a különböző szarvasfajokból izolált tetranukleotid mikroszatelliták gímszarvas ortológjainak szekvenciáját csak a lokuszok klónozásával és újra szekvenálásával lehetett megismerni, amelyhez a munkacsoportunk által kifejlesztett ún. „zoo-kloning” eljárást alkalmaztuk (Molnár *et al.*, 2007; Gyurján *et al.*, 2007; Borsy *et al.*, 2009; Stéger *et al.*, 2010). A munkánk kezdetén a GenBankban 22 tetranukleotid marker volt elérhető, amelyeket öszvérszarvasból (*Odocoileus hemionus*) valamint amerikai vapitiből (*Cervus canadensis*) izoláltak (Jones *et al.*, 2002; Meredith *et al.*, 2004). Észrevettük, hogy a publikált 25 mikroszatellita klónban 3 fordított komplementer pár volt (C02-C36, T268-T530, T40-T27), ezért összesen 22 mikroszatellitából indulhattunk ki (Szabolcsi *et al.*, 2008).

Ahhoz, hogy a gímszarvas mikroszatellita ortológ szekvenciákat felsorozhassuk, a forrásfajból származó mikroszatelliták flanking régiójára komplementer (nem

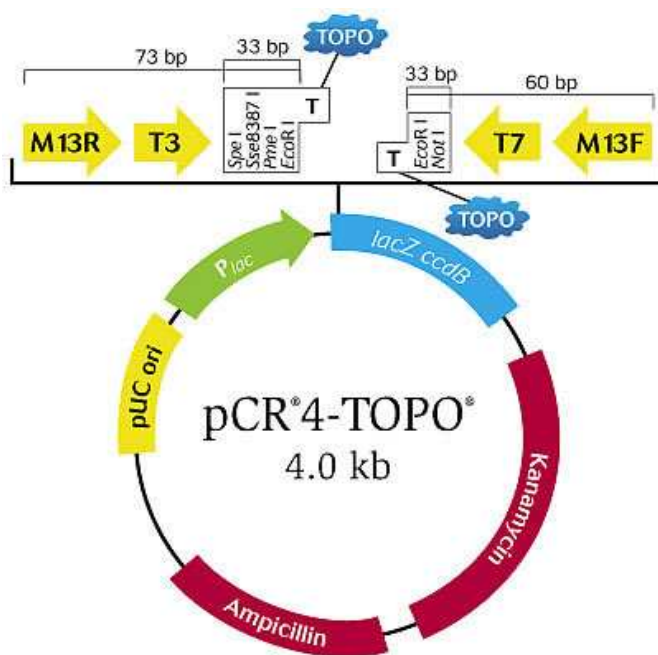
fajspecifikus) primerekkel gímszarvas genomi DNS-en PCR reakciót végeztünk. A nem fajspecifikus primereket az OLIGO ver4.0 szoftver (Rychlik, 2007) segítségével az adatbázisban lévő 22 - öszvérszarvas és vapi - mikroszatellita szekvencia 5' és 3' végeire terveztük oly módon, hogy az amplikon, az ismétlődő régió mellett minél hosszabb határoló régióval rendelkezzen. Ügyeltünk arra is, hogy az amplifikáció során az adott anneálási hőmérséklet (T_{an}) mellett primer dimerek és ún. „hairpin” struktúrák ne képződhessenek. Figyelembe vettük azt is, hogy a nem fajspecifikus primer és a gímszarvas templát szekvenciája között báziseltérések (ún. „mismatch”-ek) lehetnek. Ez a jelenség gátolhatja a primer templáthoz kötődését és ezzel együtt a PCR reakciót. Ennek kiküszöbölése érdekében az adott primer párhoz tartozó T_{an} értéket úgy választottuk meg, hogy az 3-4 °C -al a kalkulált T_m érték (ún. melting temperature – oladási hőmérséklet) alatt legyen. Az oligo szekvenciák T_m értékét a METABION internetes weboldalán elérhető szoftver segítségével számoltuk ki (<http://www.metabion.com/support/biocalc.php>). A tervezett primer a 13-50 nukleotid nagyságtartományba esik, így az összefüggés a

$$T_m = 100.5 + (41(yG + zC)/(wA + xT + yG + zC)) - (820/(wA + xT + yG + zC)) + 16.6 \log[I]$$
 egyenlet szerint számolható, ahol x , y , z és w az adott nukleotidok száma, I az egyszeres töltésű kation egyensúlyi koncentrációja, ami az alkalmazott PCR reakcióban 50 mM.

Annak érdekében, hogy az Y-kromoszómás mikroszatelliták amplifikációját elkerüljük, templátként gímszarvastehénből izolált genomi DNS-t választottunk. A PCR reakció 25 µl-es végtérfogata (a kiválasztott gímszarvastehen egyedből izolált) 1-5 ng DNS templátot, 25-25 pmol forward és reverz primert, 2,5 µl 10x PCR puffert (Applied Biosystems, 10x PCR Buffer II), 2,5 µl 25mM-os MgCL₂-ot (Applied Biosystems, 25mM MgCL₂), 1 µl 10mM dNTP keveréket és 0,5 µl AmpliTaq[®] Gold (5 Units/µl, Applied Biosystems) DNS polimeráz enzimet tartalmazott. A PCR reakció ciklusai a következők voltak; 10 min 94°C -os kezdeti denaturációt 35 ciklusban 94 °C -on 45 sec denaturáció, majd X°C -on 40 sec primer hibridizáció (anneálás) és végül 72°C -on 1 min 20 sec DNS komplementer szál szintézis (elongáció) követett. A PCR végső elongációja 72°C -on 10 min volt. A PCR termékek méretét agaróz-gélelektroforézissel vizsgáltuk. A PCR termékből 8 µl-t mértünk be a 2 %-os agaróz gél zsebeibe. A elektroforézist 1xTBE pufferben végeztük. A relatív méret meghatározáshoz standardként a 100-1000 bp DNS létrát (Fermentas GeneRulerTM) használtunk. A

primer szekvenciákat és az anneálási hőmérsékleteket (X) táblázatos formában közöltük (F.1. táblázat).

Azokat a gímszarvas lokuszokat, amelyeken a PCR reakciók eredményesnek bizonyultak, vagyis a PCR termék(ek) mérete hozzávetőlegesen megfelelt a várt – ortológ szekvencia alapján kalkulált – méretnek, a szekvencia meghatározás érdekében klónoztuk. A PCR fragmenseket pCR™ 4-TOPO® TA (Invitrogen, 1. ábra) vektorba ligáltuk, majd *E. coli* TOP10 baktériumtörzsbe transzformáltuk.



1. ábra pCR™ 4-TOPO® TA klónozó vektor és annak topoizomeráz enzimekhez (TOPO) kötött klónozó helye, valamint az M13 oligok kötőhelyei.

A ligálási reakció 6 µl-es végtérfogata 2,5µl PCR terméket, 1 µl sóoldatot (1,2M NaCl, 0,06M MgCl₂), 0,7 µl plazmid vektort (10ng/µl) és 1,8µl vizet tartalmazott. A transzformáláshoz használt TOP10 kompetens sejteket a kit előírásai szerint -80 °C-on tároltuk. Transzformálás során a sejteket 10 percig jégen tartottuk, majd a ligálási reakcióval elegyítve további 5 percig jégen hagytuk. 42 °C -on 1 perc hősokkot követően 5 percen át jégen hűtöttük a sejteket, visszafogva ezzel a baktérium sejtek saját restriktív-modifikációs rendszerét. A szelektív táptalajon való szélesztés előtt 1 órán keresztül antibiotikum-mentes 250 µl SOC tápoldatban, 37 °C -on regeneráltuk a sejteket. 120 µl sejteket tartalmazó SOC médiumot szélesztettünk 80 µg/ml ampicillin tartalmú LB agarlemezre, majd a lemezeket egy éjszakán át 37 °C -os termosztátba

helyeztük.

A klónozás sikerét a felnőtt baktériumtelepekből származó plazmid DNS klónozó helyét közrefogó primerekkel (M13) PCR reakció során ellenőriztük. Plazmid DNS tisztításához az QIAprep Spin Miniprep Kitet használtuk (Qiagen). A pozitív klónok valódiságát szekvenálással igazoltuk, amelyet ABI PRISM[®] 3100 DNS szekvenáló készüléken végeztek a Mezőgazdasági Biotechnológiai Kutatóközpont laboratóriumában. A gímszarvas nukleotid szekvencia adatok elérhetők a GenBank adatbázisban (<http://www.ncbi.nlm.nih.gov/sites/entrez?db=Nucleotide>) (F.2. táblázat).

2.2.4 Gímszarvas mikroszatelliták polimorfizmus szelekciója

Célunk az volt, hogy a marker készletünkben a legtöbb alléllal rendelkező mikroszatelliták szerepeljenek. A szakirodalomból ismert volt, hogy egy mikroszatellita lokusz közel rokon fajokból származó ortológjainak polimorfizmusa eltérő lehet (*Poetsch et al., 2001*). Ennek ismeretében nem támaszkodhattunk Jones és Meredith (*Jones et al., 2002; Meredith et al., 2004*) vizsgálati eredményeire, lokuszaink polimorfizmusát meg kellett vizsgáljuk. Mintának, az északkelet-, és délnyugat magyarországi állományból véletlenszerűen 14-14 referencia egyedet választottunk (*Szabolcsi et al.*).

A 14 gímszarvas tetranukleotid mikroszatellita szekvenciára fajspecifikus PCR primereket terveztünk az OLIGO ver4.0 szoftver segítségével. Az oligokat úgy terveztük, hogy a PCR termékek méretet hozzávetőlegesen a 100-200 bp mérettartományba essenek és a primerekben belső hurkok („hairpin”-ek), primer dimerek ne képződhessenek. A templát specificitás maximalizálása érdekében a primerek hossza meghaladta a 20 nukleotidot valamint az optimalizált anneálási hőmérséklet (T_{an}) minden primerpár esetében egyenlő volt vagy 1–2 °C -al meghaladta a primer oligok – 2.2.3 fejezetben leírt módon kalkulált – átlagolt T_m értékét.

14 referencia egyeden lokuszonként (monoplex) PCR amplifikációt végeztünk. A PCR reakció összetétele az AmpliTaq[®] Gold (Applied Biosystems) DNS polimeráz enzim végkoncentrációját (0,1 unit/μl) kivéve azonos az ortológ szekvenciák alapján kivitelezett PCR eleggyel. A reakció ciklusai a következők voltak; 10 min 94 °C -os kezdeti denaturációt 30 ciklusban 94 °C -on 45 sec újabb denaturáció, majd X °C -on 40

sec anneálás és végül 72 °C -on 1 min elongáció követett, végső elongációja 72 °C -on 10 min volt. Az amplikonok detektálását és elválasztását nagy felbontású poliakrilamid gélelektroforézissel végeztük (*Buzás and Varga, 1995; Pamjav et al., 1999; Triga et al., 1999*). Figyelembe véve azt, hogy a tetranukleotid mikroszatelliták alapvető ismétlődő egységei 4 nukleotid méretűek, az allélok elkülönítéséhez és ebből következően a lokuszok polimorfizmusának a teszteléséhez elengedhetetlen a 4 bázispárnál nagyobb felbontású gélelektroforézis. A tömörítő poliakrilamid gél 5%T 3%C, míg az elválasztó gél 7%T 7%C összetételű volt, amely a 100-120 bázispár mérettartományban akár az 1 bázispár elválasztását lehetővé tette. Az elektroforézishez a PhastSystem (Pharmacia) készüléket használtuk. A készülék a mintafelvitelt, az elválasztást és a festést automatizált módon végezte. Ennek körülményeit 1. táblázatban foglaltuk össze. A futtató puffert (1x TA) tartalmazó agaróz csíkokat – az ún. „buffer strips” – a laboratóriumban állítottuk elő. A PCR termékből 0,5 µl-t mértünk a poliakrilamid gélre.

1. táblázat PhastSystem poliakrilamid gélelektroforézis rendszer alkalmazott PCR termék futtatási és (ezüstnitrát alapú) festési folyamatai.

	<i>Fázisok</i>	<i>Hőmérséklet</i>	<i>Időtartam</i>
Futtatás	<i>10mA, 45Vh, 15°C, ~20 perc</i>		
Festés	<i>Fixáló oldat</i>	20°C	<i>5 perc</i>
	<i>Érzékenyítő oldat</i>	50°C	
	<i>Mosás -2x</i>		
	<i>Festő oldat</i>	40°C	<i>8 perc</i>
	<i>Mosás -2x</i>	30°C	<i>0,5 perc</i>
	<i>Előhívás -2x</i>		<i>0,5/4 perc</i>
	<i>Stop oldat</i>	50°C	<i>2 perc</i>
	<i>Tartósító oldat</i>		<i>3 perc</i>

A polimorfizmus teszt során a poliakrilamid gélen előhívott PCR fragmensek méretének meghatározásával elvégeztük a detektált alléltartományok becslését. A kivitelezéshez a 100-1000 bázispár DNS (Fermentas, GeneRulerTM) fragmens méretstandardot használtuk.

2.2.5 PCR fragmensek meghatározása, elválasztása és szelektálása kapilláris gélelektroforézissel

A nagyobb felbontású fragmens méret meghatározáshoz ($\pm 0,5$ nt) - vagyis a gélelektroforézisnél pontosabb genotipizáláshoz - a kapilláris elektroforézist választottuk, amelyhez az ABI PRISM[®] 3130 készülék állt rendelkezésünkre. A forward primerek 5' végét fluoreszcens festékkel jelöltük, amelyhez a DS-33 (Applied Biosystems) standard mátrix fluorofórait választottuk. A fluoreszcens oligokkal felsokszorozott DNS-szakaszok méret szerinti elválasztását kapilláris-gélelektroforézissel végeztük, amelyhez ABI PRISM[®] 3130 készüléket használtunk. Az elektroforézis körülményei a következők voltak; POP-4 polimerrel töltött 36 cm-es kapilláris, 8 sec injektálás és 15 kV feszültség mellett az elválasztás 2000 sec-ig tartott. A mintákhoz kevert belső méretstandard - GeneScan[™] 500 LIZ (Applied Biosystems) - segítségével a GeneMapper[®] ID ver3.2 szoftver (Applied Biosystems) linearizált illesztéssel kiszámította a készülék által detektált fragmensek pontos méretét. A szoftver a detektált jel alakját is vizualizálta, amelynek morfológiája alapján a mikroszatellita allélok elkülöníthetők (Butler, 2011). A kapilláris elektroforézishez a következő összetételű mixet használtuk; 1 μ l PCR termék, 0,2 μ l belső méretstandard és 9,8 μ l Hi-Di[™] formamidot (Applied Biosystems). A denaturálás ideje 95 °C -on 3 perc volt.

2.2.6 Multiplex rendszerek létrehozása

Célunk két – 5-5 mikroszatellita lokusz paralell amplifikációját egyidőben végző - multiplex PCR rendszer (két 5-Plex PCR) összeállítása volt. A multiplex PCR reakciókban a lokuszok hatékonyabb szétválasztása érdekében 5 színes elektroforézis rendszert terveztünk. Első lépésben a 10 mikroszatellita lokuszt úgy kombináltuk a két 5-Plex rendszerben, hogy kettő marker kivételével mindegyik STR külön 5' flouorórt hordozzon, valamint az azonos jelölést hordozó mikroszatelliták allélméret tartományai között a legkisebb távolság 20 nukleotid legyen. Arra is ügyeltünk, hogy az allél mérettartományok felső hátára ne haladjon meg a 250 nukleotid egységet.

Második lépésben a rendelkezésre álló primereket az OLIGO ver4.0 szoftver felhasználásával úgy módosítottuk, hogy az „erősebb” primer interakciókat (primer

dimerek), amelyek T_m értéke 50 mM $[K^+]$ koncentráció mellett 50°C feletti értéket mutatott, ne tartalmazzon. A primerpárokkal először monoplex PCR reakciót végeztünk, amelynek körülményeit az 2.2.4 fejezetben már részleteztük. A PCR termékek mennyiségét és minőségét 7,14%T 5%C összetételű poliakrilamid gél felhasználásával Multiphor II készüléken (Pharmacia) ellenőriztük. A futtatás TB-BRFK oldatban 150mA, 225Vh és 15°C volt. A fragmenseket ezüstnitrát alapú festését Budowle és munkatársai által kidolgozott módszere szerint végeztem (Budowle *et al.*, 1991). A PCR termékeket a 2.2.5 fejezetben ismertetett módon kapilláris gélelektroforézissel vizsgáltuk.

Harmadik lépésben a két 5-Plex PCR (*DeerPlex I-II*) reakciót a dNTP, Mg^{2+} , primerkoncentráció és az anneálási hőmérséklet függvényében úgy optimalizáltuk, hogy az allélikus eredetű PCR termékek mennyiségét maximalizáljuk, a lokuszok közötti imbalanszt kiegyenlítsük valamint a PCR artifaktumok (dupla „peak”, primer dimerek és nem STR termékek) mennyiségét minimalizáljuk. A templát minden esetben ugyanazon gímszarvas egyed 1-5 ng mennyiségű genomi DNS-e volt. Az optimalizált PCR reakció desztillált vízzel kiegészített 25 µl-es végtérfogata 1-5 ng DNS templátot, 1x hígítású PCR puffert (Applied Biosystems, 10x PCR Buffer II), 2,75 mM $MgCl_2$ -ot (Applied Biosystems, 25mM $MgCl_2$), 1.2 mM dNTP keveréket és 0.5 µl AmpliTaq® Gold (5 Units/µl, Applied Biosystems) DNS polimeráz enzimet és az F.2. táblázatban feltüntetett primer keveréket tartalmazott. A reakció ciklusai a következők voltak; 10 min 95 °C -os kezdeti denaturációt 29 ciklusban 94 °C -on 40 sec újabb denaturáció, majd 62/64 °C -on (*DeerPlex I* – 62 °C, *DeerPlex II* – 64 °C) 40 sec anneálás és végül 72 °C -on 1 min elongáció követett, végső elongációja 60 °C -on 60 min volt. A PCR terméket a 2.2.5 fejezetben ismertetett módon analizáltuk, minden detektált allélra meghatároztuk a számított allélméret szórását (SD) és az allélikus méret-tartományt (ún. size range).

A minták elektroferogramjának kiértékelése során a T172 marker alléljainál ún kettős csúcsot detektáltuk. A jelenség hátterében az áll, hogy a Taq polimeráz enzim a forward szál 3' végének adenilációját a PCR termékek egy részén nem fejezi be. Ennek eredményeképpen az elektroferogramon az allél csúcsok kettősek („split-peak”) vagy vállasak („shoulder-peak”) lesznek, ami megnehezíti az allélok méret szerinti azonosítását. A jelenséget részben a T501 marker alléljainál nem kettős, hanem hármas

csúcsokat detektáltunk. A jelenséget részben a részleges láncvégi adenilációnak tulajdonítottuk.

A Taq polimeráz enzim hatékonyabb templát független adenilációjának eléréséhez mind a két marker esetében az ún. „pig-tailing” eljárást alkalmaztunk, vagyis a reverz primer oligok 3' végéhez a templát DNS szállal nem komplementer GTTCTT szekvenciát szintetizáltattunk (*Brownstein et al., 1996; Butler, 2011*).

2.2.7 STR Allélok szekvencia sorrendjének meghatározása

Az STR lokuszok ismétlődő egységeinek struktúrájának összetételét szekvenálással vizsgáltuk meg. Ehhez homozigóta genotípusú egyedeket választottunk. A polimorfizmus teszthez használt primer oligo párokkal PCR amplifikációt végeztünk. A kapott PCR terméket QIAquick PCR Purification Kit-el (Qiagen) tisztítottuk. A forward és a reverz primerekkel direkt szekvenálást végeztünk a BigDye Terminator v1.1 kit felhasználásával. A reakció során kapott terméket a Performa® DTR Ultra 96-Well Plates (Edge Biosystems) segítségével tisztítottuk. A szekvencia analízist ABI PRISM 3130 készülékkel és Chromas Lite 2.0 szoftverrel végeztük. Néhány allél szekvencia vizsgálatához csak heterozigóta egyed állt rendelkezésünkre. Ezekben az esetekben a kapott PCR termékekre az 2.2.3 fejezetben ismertetett klónozási és DNS szekvenálási eljárást alkalmaztuk.

2.2.8 Populációs egyedek genotipizálása a DeerPlex I-II rendszerekkel

A *DeerPlex I-II* lokuszokkal végzett genotipizálás alapvetően két részből áll. Első lépésben az egy egyedhez tartozó két multiplex PCR reakció során felsokszorozott allélikus eredetű DNS-szakaszok méretének meghatározására került sor az ABI PRISM® 3130 kapilláris-gélelektroforézis készülék és a GeneMapper® ID ver3.2 szoftveren végzett analízis alkalmazásával (ld 2.2.5 fejezet). Mivel az STR lokuszok a hossz-polimorf markerekhez tartoznak, így a futtatás során a DNS-szakaszok méretére kapott adat az allél típusára vonatkozó információnak felelt meg. A genotipizálás második részében a mért, és morfológiája alapján allélikus eredetű csúcsokat 1 nukleotid kiterjedésű és maximálisan 0,25 nt szórású mérettartományokba osztottuk.

2.2.9 A DeerPlex I-II rendszerek fajspecifikusságának tesztelése

A multiplex PCR rendszerek fajok közötti keresztreakciójának tesztelését több párosujjú patás faj DNS-én valamint emberi eredetű DNS-en is elvégeztük ugyanis az igazságügyben vizsgálat helyszíni állatminták esetében a humán eredetű DNS-el történő keveredés lehetősége minden esetben fennáll. A fajspecifikusság tesztelése a nem-humán igazságügyi DNS vizsgálatokban előírt nemzetközi ajánlás (Linacre *et al.*, 2011). A *Ruminantia* alrendbe tartozó fajok közül azokat vontuk be a vizsgálatba, amelyek igazságügyi szempontból vagy célfajok, vagy kontaminátorok lehetnek. Vadsertésből (*Sus scrofa*), muflonból (*Ovis orientalis*), őzből (*Capreolus capreolus*) és dámszarvasból (*Dama dama*) izomszövet mintákat gyűjtöttünk, majd azokból a DNeasy Blood & Tissue Kit (QIAGEN®) felhasználásával genomi DNS-t tisztítottunk. A minták DNS tartalmát az 2.2.2 fejezetben ismertetett módon kvantifikáltuk. Szarvasmarha (*Bos taurus*) és humán DNS-hez a jelenleg is piaci forgalomban lévő kontroll DNS-ekből jutottunk hozzá, amelyeket a kvantifikáláshoz standardként használtunk (Stockmarks® Bovine Identification Kit Control DNA, AmpFISTR® Yfiler® PCR Amplification Kit Control DNA - Applied Biosystems). A fajspecifikusság teszteléséhez 5 ng templát DNS-t használtunk. A *DeerPlex I-II* multiplex PCR-ek fragmens analízisét a 2.2.5 fejezetben ismertetett módon végeztük el.

Standard oldatok:

5xTBE puffer:	445mM Tris, 445mM bórsav, 10mM EDTA (pH 8.0)
TB-BRFK oldat:	60,55g Tris, 17,31g bórsav, 2 ml 0,2% brómfenolkék/ 1000ml vízben
LB médium:	10g Bacto tripton, 5g Bacto élesztő kivonat, 10g NaCl /1000 ml vízben (pH 7.0)
LB agar lemez:	12g agar/1 liter LB médium
Tömörítő törzsoldat:	40%T 3%C; 19,4g akrilamid, 0,6g biszakrilamid /50ml vízben
Elválasztó törzsoldat:	40%T 7%C; 18,6g akrilamid, 1,4g biszakrilamid /50ml vízben
Tömörítő gél:	5%T 3%C; 0,625 ml tömörítő törzsoldat, 0,5ml 10xTA puffer,

	25µl 10% APS, 5µl TEMED /5ml vízben
Elválasztó gél:	7%T 7%C; 1,75 ml elválasztó törzsoldat, 1 ml 10xTA puffer, 50µl 10% APS, 5µl TEMED /10 ml vízben
Agaróz csík:	2,8% agaróz, 0,55% SDS, 0,2M TRIS, 0,2M Tricin
10xTA puffer:	6,78g TRIS, konc. ecetsav pH= 6,4 /50ml vízben
SEB oldat:	10mM Tris, 100mM NaCl, 39mM DTT, 10mM EDTA, 2% SDS
TE oldat:	10mM Tris, 0,1mM EDTA (pH 8.0)

2.3 Populáció genetikai és statisztikai módszerek

2.3.1 Hardy-Weinberg egyensúly (HWE) teszt

A HW egyensúly tesztelését az Arlequin v3.1 (*Excoffier et al., 2006*) és a Genepop v4.0 (*Rousset, 2008*) szoftverekkel is elvégeztük. A nullhipotézis a megfigyelt diploid genotípusokban az allélok véletlenszerű párosodása volt. A HW tesztet szeparáltan, lokuszonként végeztük el. Tesztstatisztikának mindkét esetben a „Hardy-Weinberg exakt tesztet” alkalmaztuk, amely a Fisher-féle teszthez nagyon hasonló, azzal a különbséggel, hogy a megfigyelt allélok számának megfelelő méretű ($k \times k$) kontingencia táblát alkalmaz (*Guo and Thompson, 1992*). A p -érték kiszámításához a szoftverek Markov láncsal alternatív táblákat generálnak. Minden generált $k \times k$ kontingencia táblára L_0 értéket kalkulálnak, ami annak a valószínűsége, hogy az adott $k \times k$ táblát megfigyelhetjük az allélok random asszociációját feltételezve (*Excoffier et al., 2006*). A generált $k \times k$ méretű táblákra átmeneti valószínűségi értéket $R = L_{(n+1)} / L_{(n)}$ számolnak (*Guo and Thompson, 1992*). Amennyiben R értéke 1-nél magasabb akkor az átmenet megtörténik. A módszer rendkívül érzékeny a kiindulási állapotra. Ahhoz, hogy a generált, alternatív kontingencia táblák allélfrekvencia értékei az általunk megfigyelt adatoktól minél kevésbé függjenek (minél jobban eltérjenek egymástól), a Markov láncban dememorizációs lépésekre van szükség. A lépések számát valamint a generált táblák számát a szoftverek által ajánlott értéken végeztük el (*Excoffier et al., 2006; Rousset, 2008*). A lokuszonként kapott p értékek a generált $k \times k$ méretű táblák azon hányadát adja meg, amelyeknek az L értéke a kiindulási táblával megegyező vagy

kisebb volt. Ha tehát a p érték kisebb, mint az 5 %-os konfidencia szint akkor a nullhipotézis, vagyis az allélok lokuszon belüli random párosodása (HWE) elvethető.

2.3.2 Kapcsoltsági egyensúly (LE) teszt

A mikroszatellita lokuszpárok közötti kapcsoltsági egyensúly tesztelését az Arlequin v3.1 szoftverrel végeztük el. A populációs minták több lokuszos genetikai profiljaiban a gaméták haplotípusai nem ismertek, ezért tesztstatisztikaként a valószínűségi hányados (*likelihood-ratio*) tesztet használtuk. A valószínűségi hányados $S = -2\log(L_{H^*}/L_H)$ Chi-négyzet eloszlást követ. L_{H^*} a létrejött adatok valószínűsége kapcsoltsági egyensúlyt feltételezve, vagyis a hipotézis szerint a haplotípus frekvenciák az allélgyakoriságokból multiplikálhatók. L_H az adatok valószínűsége a kapcsoltsági egyensúly hiányában. Ebben az esetben a haplotípus frekvenciák becslését az EM algoritmus segítségével végzi. A szoftver nagyon sok lépésben (esetünkben 10000) lokuszonként permutálja a meglévő allélokat, majd minden lépésben elvégzi az L_{H^*} és az L_H becslését. Ez a permutációs módszer adja az S érték nulleloszlását. Megjegyezzük, hogy a HW egyensúlytalanság a kapcsoltsági egyensúly szignifikáns eltérését okozhatja (*Excoffier and Slatkin, 1998*).

2.3.3 Bonferroni korrekció

Ha a Hardy-Weinberg és a kapcsoltsági egyensúlyt a lokuszok összességén vizsgáljuk, akkor a statisztikai elemzés során a lokuszok ugyanazon hipotézisen elvégzett teszt sorozatban vesznek részt. Ilyen esetben az általánosan használt 5%-os szignifikancia szint nem használható. Ennek az oka az alfa inflációnak („alpha inflation”) nevezett jelenség, ami egymástól független tesztek esetén matematikailag is könnyen levezethető. α_C a hiba valószínűsége a teszt során, α_E az egész vizsgálatra vonatkozó hiba valószínűsége és n darab tesztet végzünk a vizsgálat során. Annak a valószínűsége, hogy

- az egyes tesztek során nem követünk el hibát: $(1 - \alpha_C)$,
- „ n ” darab teszt során nem követünk el hibát: $(1 - \alpha_C)^n$,
- „ n ” darab teszt során legalább egyszer hibát követünk el: $1 - (1 - \alpha_C)^n$.

A fenti levezetés alapján tehát $\alpha_E = 1 - (1 - \alpha_C)^n$, ami alapján 5 teszt során annak a valószínűsége, hogy legalább egyszer hibát követünk el, már nem 5%, hanem 22,62%.

Az alfa infláció mértéke nem független tesztek - így a kapcsoltsági egyensúly tesztelése során alkalmazott páros összehasonlítások - esetén komplikáltabb. Ilyenkor csak azt tudjuk, hogy $\alpha_C \leq \alpha_E \leq n \times \alpha_C$. Ebben az esetben, ha az egész vizsgálatra vonatkozó hiba valószínűségét szeretnénk 5%-os szinten tartani, akkor az n összehasonlítás során α_C/n - nál kisebb p -értékű eredményeket kell szignifikánsnak tekintenünk. Ezt az eljárást nevezik Bonferroni-féle korrekciónak (Ewett and Weir, 1998; National Research Council, 1996; Reiczigel et al., 2010).

2.3.4 Populáció statisztikai fogalmak

A mikroszatelliták polimorfizmusának számszerűsítéséhez számos mérőszámot használtunk, amelyeket az alábbiakban ismertetünk.

Nei-féle géndiverzitás (h): egy lokuszra $h = n(1 - \sum x_i^2)/(n-1)$; a több lokuszos géndiverzitás $h_{komb} = 1 - \prod (1 - h_i)$. n a mintában előforduló allélok száma, x_i az i -dik allél megfigyelt gyakorisága (Nei, 1973; DeGiorgio et al., 2010). A számítása megegyezik a heterozigóták elfogulatlanul becsült gyakoriságával (H_E). A géndiverzitás egy olyan valószínűség, ami azt fejezi ki, hogy a populációból véletlenszerűen kiválasztott két haplotípus különbözik.

A *polimorfizmus információs tartalom* (PIC) kifejezhető annak a valószínűségeként, hogy a szülői genotípusban egyértelműen meghatározható az utódba öröklődő allél, vagy más definíciót használva egy szülő-utód relációban az anyai és az apai allélok egyértelműen meghatározhatók. Természetesen két olyan eset lehetséges ahol ez nem következtethető ki, akkor, ha az egyik szülő homozigóta és akkor, ha mindkét szülő valamint az utód is ugyanazt a heterozigóta genotípust hordozza. A PIC értéket minden

lokuszra kiszámítottuk: $PIC = 1 - \left(\sum_{i=1}^n p_i^2 \right) - \left(\sum_{i=1}^{n-1} \sum_{j=i+1}^n 2p_i^2 p_j^2 \right)$, ahol p_i az i -edik és p_j az j -edik allél megfigyelt gyakorisága, n az allélszám (Botstein et al., 1980; Guo and Elston, 1999).

Igazságügyi megkülönböztető erő (power of discrimination, PD): azt fejezi ki, hogy két egyedet a populációból véletlenszerűen kiválasztva átlagosan mekkora a valószínűsége annak, hogy a két egyed a vizsgált lokuszon eltérő genotípust hordoz. Az *átlagos egyezési valószínűség (probability of match, pM)* a PD-ből kifejezhető: $PD = 1 - \sum x_i$, több lokuszra számolva $PD_{komb} = 1 - \prod (1 - \sum x_i)$. PD-ből kalkulált egy lokuszos $pM = 1 - PD$, több lokuszra kombinálva $pM_{komb} = \prod pM_i$, ahol x_i az i -edik genotípus megfigyelt gyakorisága (Jones, 1973).

Az *átlagos egyezési valószínűség (probability of match, pM)* megfeleltethető a konzerváció genetikában használt *átlagos azonosság valószínűségének (probability of identity, PI_{ave})*. Mindkét érték annak a valószínűséget adja meg, hogy a vizsgált populációban véletlenszerűen kiválasztott két egyedben ugyanazt a genetikai profilt figyelhetjük meg. A kalkulációhoz az *API-CALC 1.0* szoftvert használtuk (Ayres and Overall, 2004), amely a Paetkau és mtsai szerint kidolgozott 10 lokuszos *PI_{ave} (pM)* érték számítását - $PI_{ave} = \prod (\sum_i p_i^4 + \sum_i \sum_{j>i} (2p_i p_j)^2)$ - az Ayres és munkatársai által kidolgozott algoritmus szerint korrigálni tudja a populáció heterogenitásával (F_{ST}) és a beltenyészet koeficienssel (F_{IS}) (Paetkau and Strobeck, 1994; Ayres and Overall, 1999).

Az igazságügyi genetikában az adott célgenotípusra alkalmazott pM számítási módszerére - a Bayes-elvű hipotézistesztesztelés alkalmazása mellett (National Research Council, 1996) - ugyancsak az Ayres-féle kalkulációt használtuk. A pM korrigált értéke az F_{ST} és az F_{IS} paraméterekkel:

$$P(A_i A_i | A_i A_i) = \frac{(F_{ST} + (1 - F_{ST})p_i)}{F_{IS} + (1 - F_{IS})(F_{ST} + (1 - F_{ST})p_i)} \left[\frac{F_{IS}^2 + 2F_{IS}(1 - F_{IS}) \frac{(2F_{ST} + (1 - F_{ST})p_i)}{(1 + F_{ST})}}{(1 + F_{ST})} + (1 - F_{IS})^2 \frac{(2F_{ST} + (1 - F_{ST})p_i)(3F_{ST} + (1 - F_{ST})p_i)}{(1 + F_{ST})(1 + 2F_{ST})} \right]$$

$$P(A_i A_j | A_i A_j) = 2(1 - F_{IS}) \frac{(F_{ST} + (1 - F_{ST})p_i)(F_{ST} + (1 - F_{ST})p_j)}{(1 + F_{ST})(1 + 2F_{ST})}, \text{ ahol}$$

$P(A_i A_i | A_i A_i)$ a véletlenszerű egyezés valószínűsége homozigóta célgenotípus esetén,
és $P(A_i A_j | A_i A_j)$ a véletlenszerű egyezés valószínűsége heterozigóta célgenotípus esetén.

A kis mintaszám következtében (<100) a populációs minta nem minden esetben reprezentálja a valós populációban előforduló ritka allélokat. A 0,01 %-os értéknél nem nagyobb gyakoriságú allél ritka allélnak tekinthető (Chakraborty, 1992). Az igazságügyi vizsgálatok során olyan célprofilok is előfordulhatnak, amelyekben egy vagy több allél – azok alacsony valós populációs frekvenciája következtében – nem szerepelnek a populációs adatbázisban. A minimális allélgyakorisági érték (p_{min}) bevezetése lehetővé teszi, hogy azoknak a ritka alléloknak a gyakoriságát is megbecsülhessük, amelyek a populációs adatbázisban nem szerepelnek (Budowle, 1996). A minimális allélgyakoriság egy olyan küszöbértéknek is tekinthető, amely fölött a megfigyelt allélok gyakorisága bizonyos (általában 95%-os) konfidenciával megbízhatónak tekinthető. A p_{min} értéke a lokusz heterozigóta gyakoriságának növekedésével nő, a populációs minta méretének növekedésével pedig csökken (Chakraborty, 1992):

$$p_{min} = 1 - \left[1 - (1 - \alpha)^{\frac{1}{c}} \right]^{2n} \text{ ahol } \alpha \text{ a választott szignifikancia szint, } n \text{ a mintapopuláció}$$

egyedszáma és c a gyakori allélok száma. A c érték kifejezhető a megfigyelt heterozigóta gyakoriságból (H_o) a következő összefüggés szerint:

$$T = \frac{H_o}{1 - H_o}, K = \sum_{i=0}^{2n-1} \frac{T}{T+i}, A = \sum_{i=1}^{2nr} \frac{1}{i}, B = \sum_{i=1}^{2nr} \frac{1}{2n-i}, P = AT - BT^2, c = K - P, r \text{ a ritka}$$

allélok gyakorisága. Nelson és munkatársai (Nelson, 1978) nyomán a minimális allélgyakoriság a $p_{min} = 1 - \alpha^{\frac{1}{2n}}$ egyenlet szerint is kifejezhető.

A genotípus gyakoriság konfidencia intervallumát Chakraborty és mstai által kidolgozott módszere alapján becsültük (Chakraborty et al., 1993). A mintavételi variancia egy adott (l) lokusz homozigóta és heterozigóta genotípusa esetében

$$V(P_{hom}) = \frac{4p_l^2(1-p_l)}{2n}, V(P_{het}) = 4p_lq_l \left[\frac{p_l + q_l - 4p_lq_l}{2n} \right]. \text{ Ha a lokuszok függetlenségét}$$

feltételezzük, akkor több lokuszos genotípus mintavételi varianciája a markerek szórásnégyzetének összegével egyenlő, vagyis $V[\ln(P_{komb})] \approx \sum_{l=1}^L \frac{V(P_l)}{P_l^2} = s_{komb}^2$ ahol

$$\ln(P_{komb}) = \sum_{l=1}^L \ln(P_l). L \text{ a mikroszatellita lokuszok számával, } n \text{ a minta egyedszámával}$$

egyenlő. A többlokuszos genetikai profil 95%-os biztonsággal becsült konfidencia intervallumát az $\ln(P_{komb}) \pm Z_{\alpha} s_{komb}$ összefüggés alapján definiálhatjuk, ahol Z_{α} a standard normális eloszlás kétoldali kritikus értéke, ami 5%-os hiba ($\alpha = 0,05$) esetén 1,96.

2.3.5 A populáció genetikai struktúrája

F-statisztika és molekuláris variancia analízis

Az egyes populációk közötti genetikai különbség - vagyis a több lokuszos genotípusok (genetikai profilok) populációk közötti variabilitásának - becsléséhez Wright-féle F-statisztikai analízist végeztünk (Wright, 1951). A Wright-féle beltenyészet paraméterek (F_{IS} , F_{IT} , F_{ST} ($\sim \theta$)) becslését molekuláris varianciaanalízissel (ANOVA és AMOVA) végeztük el, amelyhez a Genepop és az Arlequin szoftvert használtuk (Excoffier et al., 1992, 2006; Roussel, 2008). Mindkét módszer alapja az, hogy a teljes molekuláris variancia (σ_T^2) un. „variancia komponensekre” bontható ($\sigma_a^2, \sigma_b^2, \sigma_c^2$). Figyelembe véve azt, hogy a vizsgálatunkba bevont populációs minta két alcsoportra bontható (EK-, DNY-i alpopulációs minta), ezért σ_T^2 -nek három komponensét vizsgálhattuk. A σ_a^2 - a populációk között és a σ_b^2 - a populáción belül az egyedek között létrejött allélvariancia és σ_c^2 - az egyeden belül a gaméták között fennálló allélvariancia. A „variancia komponensek” segítségével az F-statisztikai paraméterek a következő egyenlet szerint kifejezhetők:

$$F_{ST} = \frac{\sigma_a^2}{\sigma_T^2}, \quad F_{IT} = \frac{\sigma_a^2 + \sigma_b^2}{\sigma_T^2}, \quad F_{IS} = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_c^2}.$$

Az ANOVA és az AMOVA elemzése során kapott F_{ST} azt fejezi ki, hogy a populációkban meglévő totális molekuláris variancia mekkora része tulajdonítható a populációk közötti varianciának. A kalkulált F_{IS} érték a teljes allélvarianciának azt a hányadát fejezi ki, amely a populációkon belül az egyedek között mérhető. Az F értékek statisztikai szignifikanciáját nem paraméteres permutációs teszttel határoztuk meg (Excoffier et al., 1992).

A szimulációt végző szoftver az egyedekben megfigyelhető lehetséges haplotípusokat/gamétákat véletlenszerűen újra szétosztja a teljes populáción belül (σ_T^2) és a két populáción (EK-DNY) belül (σ_b^2), valamint az egyedek genotípusait véletlenszerűen újrarendezi a két populáció között (σ_a^2). A szoftver a beállítás szerint több mint 1000 ilyen replikációt hajt végre és minden egyes lépésben F értékeket kalkulál. A szimulációval kapott p érték annak a gyakoriságát adja meg, hogy a replikációk hányad részében haladta meg a szimulációs F érték az aktuális F értéket. Ha a kapott $p < 0,05$, akkor $\alpha = 0,05$ szignifikancia-szint mellett a nullhipotézist ($H_0 : F_{IS} = 0, F_{IT} = 0, F_{ST} = 0$) elvethetjük.

„Assignment teszt”

Igazságügyi genetikában a DNS bizonyíték statisztikai interpretációja rendkívüli jelentőséggel bír. A helyszíni minta genetikai profiljával (célprofil) véletlenszerűen egyezést adó gímszarvas egyed valószínűségét befolyásolhatja a választott referencia adatbázis, ugyanis az EK és DNY alpopulációs adatbázisok hatással lehetnek a genetikai profilok gyakoriságára, ezáltal befolyásolva a DNS bizonyíték erejét. Ennek érdekében megvizsgáltuk az adatbázisok genetikai profilokra gyakorolt hatását. Az ún. „assignment teszt” (*Paetkau et al., 1995*) alkalmazásával azt is megvizsgáltuk, hogy ennek a hatásnak a mértéke prediktálhatóvá teszi-e egy gímszarvas egyed populációs eredetét.

Az adatbázis profilgyakoriságra gyakorolt hatása a becsült beltenyészet paraméteren kívül annak inhomogenitásától is függ. Az EK és DNY populációk genetikai struktúrájának analízisét az azt alkotó egyedek vizsgálatával folytattuk. A populáció homogénnek tekinthető-e, vagyis a populáción belül csoportokba rendezhetők.

Első lépésben a két populáció egyedeit egyenként kiemeltük saját (referencia) populációjukból és a saját alléljaik hiányában kiszámítottuk profil gyakoriságukat a referencia adatbázisukban, majd visszahelyeztük őket. Ezzel a lépéssel kiküszöböltük annak lehetőségét, hogy a saját adatbázis elfogult legyen az adott egyed mikroszatellita allél-együttesének gyakoriságára nézve (*Brenner, 1998*). A kiemelt egyedek profil gyakoriságát a nem referencia adatbázis allélfrekvencia értékeivel is kiszámoltuk. Ezzel az analízissel egy olyan bűnügyi helyzetet szimuláltuk, amikor pl. az illegálisan elejtett

gímszarvasból származó helyszíni vérmintából és az állat agancs trófeájából ugyanazt a 9 vagy 10 lokuszos genotípust mutatjuk ki, azaz a bűncselekmény során keletkezett biológiai nyom egy gímszarvas egyedtől származhatott. Ebben az esetben – és ha a laboratóriumi hiba kiküszöbölhető - az egyezési valószínűség (pM) megegyezik a profilgyakoriság (P) értékkel (Balding, 1995; National Research Council, 1996; Ewett and Weir, 1998). A P értéket – a kapcsoltsági egyensúly feltételezésével, a termék szabálynak megfelelően – az Ayres és Overall által kidolgozott módszerrel kalkuláltuk ki (Ayres és Overall, 1999). A populációs adatbázisainkban csak azokkal az allél frekvencia értékekkel kalkuláltunk, amelyek elérték, vagy meghaladták a p_{min} értéket. A P érték torzítatlan becslése érdekében Paetkau és munkatársai egyszer, Balding és munkatársai pedig kétszer adták hozzá a populációhoz a vizsgált egyed alléljait (Balding, 1995; Paetkau et al., 1995). Tekintetbe véve az EK-DNY populációk kis (<100 egyed) minta nagyságát, és a populációnként becsült viszonylag magas minimális allélfrekvencia értékeket a Balding és a Paetkau módszer szerint alkalmazott *posterior* allélgyakorisági értékek meghatározását nem alkalmaztuk, ugyanis csak nagy populációs adatbázisok esetén adnának elfogulatlan becslést, ahol a p_{min} értékek rendkívül alacsonyak (Balding, 1995).

Minden kiemelt egyed mindkét populáció alapján kapott P értékét sörét diagramon („scatter plot”) ábrázoltuk, amelynek x és y - tengelyén az EK és DNY adatbázis által kapott 10 lokuszos gyakorisági értékek logaritmusai szerepelnek. A populációs adatbázisonként kalkulált P értékpárok között jelentős különbség lehet, amelynek számszerűsítéséhez egy P értékpár diagramot is készítettünk. Ennek y - tengelye az egyedekhez tartozó P értékpárok tízes alapú logaritmusa, az x - tengelyén az összes, a populációs adatbázisba vett egyedet feltüntettük. A diagramon az északkeleti- (EK), délnyugati- (DNY) és teljes (T) populációs adatbázisok alapján kalkulált P értékek 95%-os konfidencia intervallumát is ábrázoltuk.

Az „assignment teszt” alapelve az, hogy az egyedek genotípus gyakorisági értéke a forráspopulációjukban a legmagasabb, vagyis az egyedek abba a populációba sorolhatók, amelyekben a legnagyobb a genotípusuk valószínűsége (Paetkau et al., 1995). A vizsgált egyed valószínűség alapú (LBM - „likelihood based method”) populációs eredetének meghatározása Bayes elvű módszert használ, vagyis egy

valószínűségi hányadost (LR) kalkulál. Az $LR_{LBM} = \frac{P_A}{P_B}$, ahol P - az egyed a több lokuszos genotípus gyakorisága az A vagy B populációban (*Yang et al., 2005*).

3 Eredmények és kiértékelésük

Molekuláris genetikai eredmények

A Genbank (<http://www.ncbi.nlm.nih.gov/genbank>) DNS szenvencia adatbázisában jelenlévő összesen 22 tetranukleotid mikroszatellita lokusz faji eredete eltérő (*Cervus canadensis* és *Odocoileus hemionus*), nem a gímszarvasból izolált STR lokuszok. Egy adott mikroszatellita ortológ határoló régióiban fajok/alfajok, de akár elkülönült populációk vagy egyedek között is szekvencia eltérések lehetnek. Ennek következtében a tervezett primer és a templát közötti esetlegesen nem teljes szekvencia komplementaritás (ún. „mismatch”) miatt a polimeráz láncreakció során allélkiesés vagy akár teljes reakció gátlás is felléphet. A PCR primerek tervezéséhez meg kellett ismerjünk a rendelkezésre álló tetranukleotid mikroszatellita lokuszok gímszarvas ortológjainak határoló (ún. „flanking”) szekvenciáit.

3.1 A „zoo-cloning” alkalmazása a gímszarvas mikroszatellita markerek azonosításához és szekvencia analíziséhez

Munkacsoportunk jelentős eredményeket ért el a gímszarvas agancs (csont) fejlődésében szerepet játszó gének kutatásában (ld 2.2.3 fejezet). A gének ismeretében sok kérdést tisztáztunk a porc-, az enchondrális csontfejlődés és a gímszarvas ciklikus fiziológiás csontitkulásában szerepet játszó génregulációban. Figyelembe véve azt a tényt, hogy a csoport tevékenységének kezdetén nagyon kisszámú szarvas cDNS szerepelt az adatbázisokban, a szarvas ortológ gének azonosításához egy ún. „zoo-klónozási” eljárásnak elnevezett módszert fejlesztettük ki. Az alapelv az, hogy a szarvasfélék (*Cervidae*) és a szarvasmarha félék (*Bovidae*) közel rokon családok, vagyis az ortológ génekben magas fokú szekvencia hasonlóságot feltételeztünk. A gímszarvas csontszövetből (borda) nyert cDNS-ek szekvencia ismerete alapján szarvasmarha/juh BLAST szekvencia adatbázisban azonosítottuk az ortológ célgént. A gímszarvas gén teljes szekvenciájának megismerése érdekében a PCR amplifikációhoz szükséges oligo

szekvenciákat a szarvasmarha vagy juh ortológ gének exon régiójára terveztük. A kapott PCR termékek szekvenálását a gímszarvas mikroszatellita géneknél ismertetett módon végeztem el. Munkám során a ciklikus fiziológiás oszteoporózisban upregulált *ENO1* gén azonosítását és szekvenciájának megismerését végeztem el (Borsy *et al.*, 2009, GenBank acc.no.: EF619484), ennek részletesebb ismertetését a dolgozat terjedelme miatt nem közlöm. A tetranukleotid mikroszatellita gének ortológjainak vizsgálata a „zoo-klónozási” eljárás alapján.

A munkánk kezdetén rendelkezésünkre álló összesen 22 tetranukleotid mikroszatellita ortológból 14 -et sikerült a gímszarvas tehén DNS temláton felsokszorozni és sikeresen plazmid vektorba klónozni, majd szekvencia sorrendjét meghatározni (Szabolcsi *et al.*, 2008). Ezeket a lokusz adatokat a Genbank szekvencia adatbázisba töltöttük fel (F.2. táblázat). Azt a 8 STR ortológot, amelyeknél vagy a nem fajspecifikus primerekkel elvégzett PCR reakció vagy a klónozás eredménytelennek bizonyult, tovább nem vizsgáltuk. A rendelkezésünkre álló 14 gímszarvas mikroszatellita ismétlődő (ún. „flanking”) régióinak szekvencia sorrendjét összehasonlítottuk a génbankban publikált *Cervus canadensis* és *Odocoileus hemionus* ortológjaival. A szekvencia illesztéseket az F.3. táblázatban közöltük.

A határoló (ún. flanking) régiókból a teljes lokusz amplifikálásához használt PCR primer oligok (ún. „klónozó primerek”) kötő helyeit kivettük az összehasonlításból, mivel azok nem a gímszarvas szekvencia templát alapján készültek. A T172 és a T123 lokuszok esetében egyik flanking régióban sem, míg a T108 lokusznál csak egy allélnál mutattunk ki szekvencia eltérést. A vapitiből izolált T107, T156, T193, T501 és a T507 markerek ortológ szekvencia illesztésénél a relatív eltérések száma (mutált nukleotid pozíció/összes nukleotidok száma) alacsony volt (T107- 1/177,50; T156- 1/90,50; T193- 1/301,00; T501- 1/74,75, T507- 1/127,66). Hasonlóan kevés bár viszonylagosan több szekvencia eltérést mutattunk ki az öszvérszarvasból izolált T180 lokusznál (relatív eltérések száma: 1/64,30), amely nagyságrendileg megfelel a C02 vapiti eredetű lokusznál detektált értéknek (relatív eltérések száma: 1/50,00). A flanking régió méretéhez viszonyított legtöbb mutációt az öszvérszarvas eredetű C01, C143, C229 valamint - nem várt eredményként - a vapiti eredetű T26 STR markerek esetében mutattuk ki (relatív eltérések száma: 1/20,00; 1/30,60; 1/30,44; 1/22,11)

A gímszarvas templát szekvencia alapján tervezett 20 ún. specifikus PCR oligo közül 7 primer kötőhelye hordozott mutációt az ortológ szekvenciákhoz viszonyítva. Ezek közül a C01 STR marker reverz primerének kötő helyében két eltérést – egy tranzíciót és egy inszerció - is kimutattunk, a többi 6 PCR oligo esetében 1-1 bázispárt érintő mutációkat azonosítottunk (4 tranzíció, egy transzverzió és egy delécio). Már egy tranzíciós komplementaritás különbség is megakadályozhatja a PCR primer anneálódását és ezzel a PCR termék létrejöttét (*Leibelt and Budowle, 2003; Takayama et al., 2007*).

3.2 Gímszarvas tetranukleotid mikroszatellita lokuszok szelektálása a polimorfizmus teszt alapján

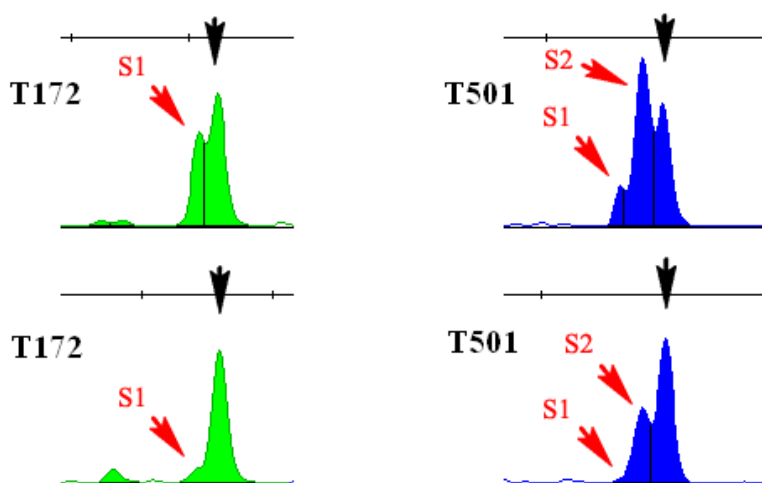
A 14 - mostmár – gímszarvas mikroszatellita lokuszból lehetett kiválogatni azt a 10 STR markert, amelyek majd a *DeerPlex I-II* multiplex PCR rendszereket alkotják. A mikroszatellita ortológok polimorfizmusa különböző fajokban igen eltérő lehet (*Poetsch et al., 2001*). Ennek ismeretében gímszarvas lokuszaink polimorfizmusát tesztelnünk kellett. A tesztet a 2.2.4 fejezetben ismertetett poliakrilamid-gélelektroforézissel végeztük. A 14 gímszarvas STR lokuszból a polimorfizmus teszt során azokat szelektáltuk ki, amelyek a 14 referencia egyedben legalább három alléllal rendelkeztek (F.2. táblázat). 10 mikroszatellita marker bizonyult polimorfnak, vagyis ezeknél a lokuszoknál a 14 egyed vizsgálata során több mint három allélt tudtunk elkülöníteni. A többi STR markert (C02, C143, C180, T107) a rendkívül alacsony fokú allél polimorfizmus (3 vagy annál kevesebb allél) miatt kihagytuk a *DeerPlex* fejlesztésének további lépéseiből (*Szabolcsi et al.*).

3.3 A *DeerPlex I-II* multiplex PCR rendszerek

A 2.2.6 fejezetben tárgyalt, a T172 marker alléljainál detektált ún. kettős csúcs kiküszöbölése érdekében a reverz primerek 3' végéhez GTTCTT oligonukleotidot szintetizáltattunk (ún. „pig-tailing” eljárás, *Brownstein et al., 1996; Butler, 2011*). A primer modifikáció hatását 23 db véletlenszerűen kiválasztott allél esetében vizsgáltuk. A PCR oligok módosítása előtt a jelenség minden vizsgált allél esetében megfigyelhető volt. A csúcsok átlagos távolsága 0,755 bázis (SD = 0,039), a két csúcsmagasság

arányának átlaga pedig 0,736 (SD = 0,095) volt. Az ún. „pig-tailing” eljárás hatására a részleges láncvégi adeniláció jelenségét a T172 lokusz esetében oly mértékben sikerült lecsökkenteni, hogy az elektroferogramok kiértékelése során az ún. kettős csúcs egy esetben sem volt jelen (2. ábra).

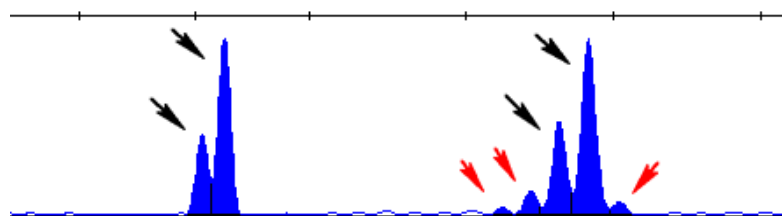
2. ábra Az ábrán felül a T172 és a T501 alléljainak elektroferogramján detektált ún. „kettős” és „hármás” csúcsok láthatóak, az ábrán alul a két marker reverz primer oligojának modifikációja („pig-tailing”) után kimutatott allélok elektroferogramja látható. Piros nyíl - „split-” vagy „shoulder-peak”-ek (S1,S2), Fekete nyíl – allélikus eredetű adenilált csúcsok (allél „peak”).



A T501 marker alléljainál a POP-4 polimerrel végzett fragmens analízis során ún. hármás csúcsot figyeltünk meg, ami megnehezítette az allélok pontos méret meghatározását (2. ábra). A „split” csúcsok (S1, S2) egymástól mért átlagos távolsága 0,952 (SD = 0,049) nt, az S2 „split” csúcs és az adenilált allél csúcs között mért átlagos távolság 0,857 (SD = 0,047) nt volt. Az S2 „split-peak” csúcsmagassága bizonyult a legnagyobb, az S1 intenzitása átlagosan 23,2 százaléka (SD = 0,05) volt az S2 csúcsnak, míg az adenilált allél intenzitása átlagosan 72,2 százaléka (SD = 0,1) volt S2-nek. A T501 marker PCR terméke mononukleotid ismétlődést tartalmaz (F.4-b. táblázat). A tíz adenin monomer ismétlődés a Taq polimeráz enzim csúszását idézheti elő minden egyes amplifikációs lépés láncnövekedési fázisában. Ennek eredménye az ún. „stutter” termék lesz (Walsh *et al.*, 1996). A Taq polimeráz enzim a monomer

ismétlődő egység felsokszorozása során nemcsak egy, hanem több „stutter” terméket is létrehozhat. Ezek a valós allél méretéhez képest akár több ismétlődő egységgel is kisebbek lehetnek (Clarke *et al.*, 2001; Fazekas *et al.*, 2010). Az ún. hármas csúcs jelenlétét részben a polimeráz enzim részleges adenilációjának részben a „stutter” termékeknek tulajdonítottuk ezért – hasonlóan a T172 markerhez – itt is az ún. „pig-tailing” eljárást alkalmaztuk. A primer módosítás hatására az elektroferogramokon az S1 csúcsmagassága oly mértékben csökkent, hogy véletlenszerűen kiválasztott 23 allélból mindössze háromszor figyeltük meg és S1 csúcsmagassága az adenilált allél értékének már csak 6,4 százaléka volt. A primer oligo módosítása után S2 ún. shoulder-peak” formájában volt kimutatható (2. ábra). A vállak átlagos távolságára 1,033 nt (SD = 0,054), a csúcsmagasság arányának átlagára pedig 0,471 nt (SD = 0,071) mértünk. A T501 lokusz módosított primerrel felsokszorozott fragmenseinek méret meghatározását POP-7 polimerrel, 50 cm-es kapillárison is elvégeztük. A nagyobb felbontás hatására az elektroferogramon a monomer ismétlődések okozta „stutter” termékek csúcsai jól azonosíthatók (3. ábra), amelyek között a valós allél méretéhez képest 3-4 nukleotiddal kisebbek is azonosíthatók.

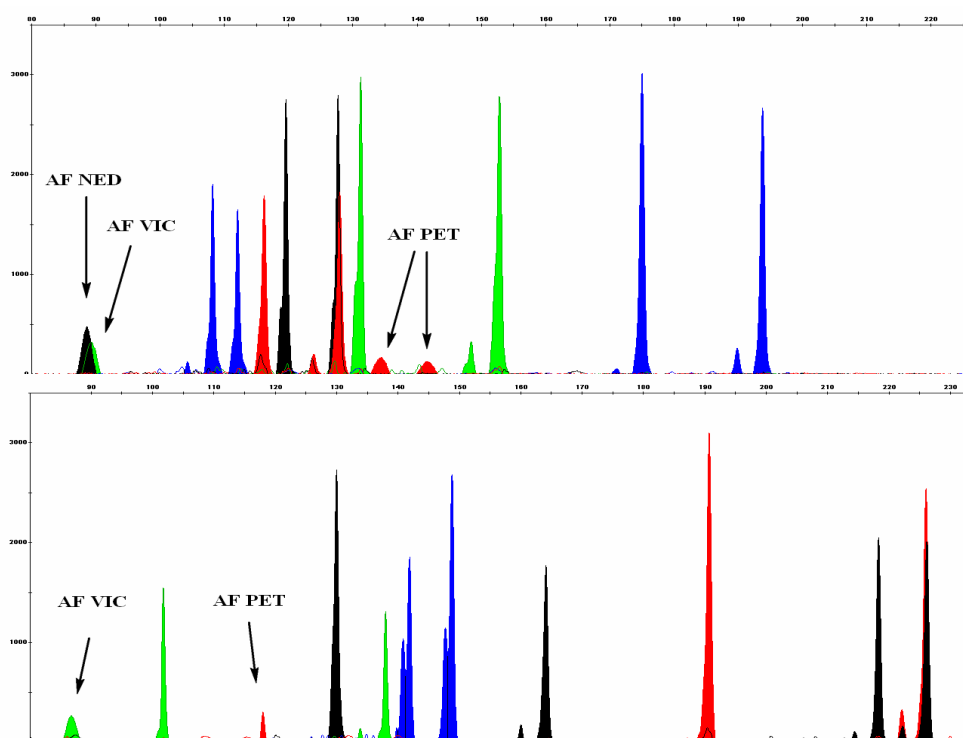
3. ábra Az ábrán bal oldalon a T501 lokusz 6.1 alléljának POP-4 polimerrel, 36 cm-es kapillárisal és a jobb oldalon a POP-7 polimerrel, 50-cm-es kapillárisal készült elektroferogramja látható. Mindkét rendszerrel detektált un. „kettős” csúcsok a fekete nyilakkal jelölve láthatóak, a piros nyilak a nagyobb felbontású elválasztás során detektálható „stutter” csúcsokat jelölik.



A két multiplex PCR reakció összetevőinek (2.2.6 fejezetben közölt) optimális arányai mellett minden allélnál egy allélikus és egy „stutter” csúcsot detektáltunk, a T501 lokusz alléljainál vállas, vagyis kettős csúcsokat.

A polimeráz láncreakció során nemcsak allélikus eredetű fragmensek hanem ún. nem specifikus műtermékek is képződhetnek. A PCR során ugyanis a genomban a véletlenszerűen anneálódó primer szekvenciák – ún. „mispriming” – egymással (forward-reverz) vagy önmagukkal felsokszorozott termékeket hozhatnak létre. Ezek könnyen felismerhetők atipikus csúcsmorfológiájukról (*Gill et al., 1997b*), és általában nem allélikus pozícióban illetve a lokuszok allélméret tartományán kívül helyezkednek el. A *DeerPlex I* PET fluoreszcens festék jeleiben két nem specifikus csúcs volt felismerhető (133,82 – SD = 0,527; 141,02 – SD = 0,51), amelyek ugyan az allélméret tartományon belül voltak, de nem estek egyik allélikus pozícióba sem. A T193 és a T108 markereknél azonos méretű, allélméret tartományon kívüli műterméket azonosítottunk (88,9 – SD = 0,364; 88,25 – SD = 0,388). A *DeerPlex II* multiplex rendszer két fluoreszcens jelében kaptunk atipikus, nem allélikus pozícióba eső 5 százalékos intenzitás meghaladó műterméket (T172(VIC): 86,34 – SD = 0,449; T507(PET): 118,01 – SD = 0,03). Ezeknek, a mispriming jelenség eredményezte csúcsoknak az intenzitása az optimális templát DNS mennyiségét alkalmazva a 100 *DeerPlex I-II* profilban egyszer sem érte el az allélikus csúcsok 15 százalékát (4. ábra).

4. ábra Az ábrán a *DeerPlex I* (felül) és *II* (alul) elektroferogramja látható. Fekete nyilakkal a műtermékeket („artifatumokat” – AF) jelöltük. AF NED,-PET,-VIC: a NED,PET és VIC fluoreszcens festékeket hordozó primerek által felsokszorozott műtermékek. Az ábra *x* -tengelyén a kalkulált fragmensméretek, az *y* -tengelyen a relatív fluoreszcencia egységek értékei láthatók. A nyilakkal nem jelölt csúcsok allélikus eredetűek.



3.4 Allélpolimorfizmusok felmérése a *DeerPlex I-II* rendszerekkel

Az anyagok és módszerek 2.2.6 fejezetében ismertetett DNS templát mennyiséggel és multiplex PCR protokoll alkalmazásával mértük fel az STR lokuszok allélpolimorfizmusát a magyarországi gímszarvas állományból származó 100 egyed alapján. A Genemapper szoftver az allélok futási idejéből és a GeneScan™ 500 LIZ (Applied Biosystems) belső standard fragmensei által illesztett kalibrációs egyenesből, lineáris illesztéssel számítja ki az allélok méretét. A detektált allélcsúcsok maximálisan $\pm 0,5$ nt kiterjedésű és maximálisan 0,25 nt szórású fragmensméret tartományokba

voltak besorolhatók. Az allélcúcsok közül a legszélesebb mérettartományt a T156 marker 9-es alléljánál mértük (0,95 -14 db allél). Három marker – T156, T172 és T26 – esetében 0,25 nukleotidnál nagyobb szórású fragmens mérettartományokat is detektáltunk (F.5-a., F.5-b. táblázatok). A kapilláris elektroforézis körülményeinek (pl. hőmérséklet, feszültségváltozás, elektroforézis puffer koncentráció stb.) változása befolyással van a - GeneScanTM 500 LIZ - méret standard és az ampliton fragmensek közötti mobilitás különbségre (*Buel et al., 2003*), vagyis a standard alapján kalkulált méretek változhatnak. A három marker bizonyos allélfragmens mérettartományának szórás növekedése ezzel az anomáliával is magyarázható. A nemzetközi szakirodalomból ismert jelenség, hogy az azonos méretű, de eltérő szekvenciájú DNS fragmensek elektroferotikus mobilitása eltérő lehet (*Butler, 2011*). A humán SE33 lokusz egyes alléljainál a vad típushoz képest egy bázispárnyi szekvencia különbség olyan másodlagos struktúrális változást eredményezett, amely 0,84 nukleotid egységgel növelte a mutáns allél átlagos méretét (*Wang et al., 2012*). A három gímszarvas lokusz egy-egy alléljánál (T156 – 9, 172 – 12, T26 – 9.1) tapasztalt 0,25 nukleotidnál nagyobb szórásértékét tehát nemcsak az elektroforézis körülményeinek változása, hanem a szekvenciavariáns allélok megléte is okozhatta. A 100 egyed allélcúcsainak méretcsoportokba rendezése során összesen 135 allélt különítettünk el (*Szabolcsi et al.*). A mikroszatelliták hossz-polimorfizmus alapú markerek, amelyek alléljait a méret és a szekvencia egyaránt jellemez. Tehát meg kellett vizsgálnunk azt, hogy az allélok milyen DNS szekvenciával és ismétlődő (repeat) egységekkel jellemezhetők. Ennek érdekében 50 allélikus pozícióból egy-egy fragmenst kiválasztva meghatároztuk azok DNS szekvenciáját. A T26, T501, T507 és a T156 lokuszokon belül olyan allélcúcs tartományokat is kimutattunk, amelyek átlagai között egy, két vagy három nukleotid különbséget mértünk. A szakirodalomból ismert tény, hogy a tetranukleotid STR markerek intermedier alléljai közötti szekvencia különbség akár egy bázispár is lehet (*Butler, 2011*), ennek lehetőségét a fent említett három marker esetében szekvenciaanalízissel meg vizsgáltuk.

3.5 A kiválasztott fragmensek szekvencia-analízise

A kiválasztott fragmenseket a kiválasztott egyedből monoplex PCR amplifikációval felsokszoroztuk. Elsősorban a homozigóta egyedeket választottuk ki, de néhány fragmens esetében csak heterozigóta egyedekre támaszkodhattunk. A homozigóta egyedekből a PCR termék tisztítása után direkt szekvenálást végeztünk, amelyekhez a markerek specifikus primereit használtuk. A heterozigóta egyedek esetében a PCR terméket először pCR4 TOPO-TA vektorba ligáltuk, majd a plazmidot *E. coli DH5 α* kompetens sejtbe klónoztuk. A pozitív klónok plazmid inszertjein M13 oligo segítségével szekvencia analízist végeztünk. Minden esetben megkaptuk az adott lokuszon heterozigóta egyed mindkét alléljának DNS szekvenciáját. A szekvenált allélokat táblázatos formában közöltük (ld F.4-a., F.4-b. táblázatok). A fragmens analízissel elkülönített 135-ből 50 allél teljes DNS szekvenciájának vizsgálatát végeztük el. A populációs felmérés során elkülönített allélokból 28 intermedier (mikrovariáns) típusú allélnak bizonyult. Ez a magas mikrovariáns arány a markerek ismétlődő régiójának összetételéből adódhatott. Az allélok közötti 1 nukleotid méretkülönbség a T507 marker 10 és 10.1 alléljai esetében egyértelműen igazolható volt. Az allélcúcsok méretének átlagai között mért átlagosan két és három bázis különbség szekvencia vizsgálattal is bizonyítható volt a T156 és a T501 markerek intermedier típusú alléljainál.

3.6 A populáció gímszarvas egyedeinek genotipizálása *DeerPlex I-II* rendszerekkel

A 10 lokuszos genotípus meghatározását az allélok mérete alapján végeztük el. A populációs egyedek allél polimorfizmus vizsgálata során mért elektroforetikus mérettartományba eső allélokat a *bevezető fejezetben* említett nemzetközi ajánlásokat alkalmazva - az ismétlődő régiók struktúrája alapján - neveztük el. A 100 egyed genotipizálása során azonos genotípust nem találtunk, az egyedek 10 lokuszos genotípusát (genetikai profilját) táblázatos formában közöltük. (F.6. táblázat)

3.7 A 10 mikroszatellita lokusz jellemzése

A 10 markerből a T108, T156, T172 és a T507 tartalmazott egyszerű ismétlődő egységeket (F.4-a., F.4-b. táblázatok). A nemzetközi ajánlások alapján az allélokat számmal jelöltük, amely az egyszerű markereknél a tetramer ismétlődő egységek számával egyenlő. A T108, T156, és a T507 markerek esetében a polimorf repeat egység a GATA, a T172 esetében - ennek reverz komplementer szekvenciája - a TATC motívum volt. A T507 marker esetében még egy addicionális (TA) dimer ismétlődést is azonosítottunk, amely a populációs vizsgálatok során nem bizonyult polimorfnak. Eltérve a nemzetközi ajánlásoktól a kettős egységeket nem számítottuk bele az allélok elnevezésébe, ugyanis ez rendkívül megnövelte volna az intermedier jelölésű allélok számát. A szekvencia analízis során kiderült, hogy a T156 markernél a (GT) dimer egység deléciója okozta az N.2 variáns allélok jelenlétét.

A C229, T123 és a T193 mikroszatelliták ismétlődő régiói összetett struktúrájúak (F.4-a., F.4-b. táblázatok). A nemzetközi ajánlások szerint az összetett markerek alléljainak elnevezésében minden azonosított egységtípust számításba kell venni. A T123 mikroszatellita (TAGA)₃ egysége – a 100 egyed vizsgálata során - nem bizonyult polimorfnak és nem szomszédos a polimorf tetramer egységekkel, ezért az allélok elnevezésekor a három (TAGA) struktúra nukleotidjainak számával nem növeltük a jelöléseket.

Az ismétlődő régió szekvencia analízise alapján a C01, T26 és a T501 mikroszatelliták komplex struktúrájúak, vagyis különböző méretű és szekvenciájú ismétlődő egységeket tartalmaznak (F.4-a., F.4-b. táblázatok). A C01 marker különböző ismétlődő egységeiből csak a CATC repeat mutatott hossz-polimorfizmust, azonban tekintetbe véve a repeat egységek szekvencia változatosságát és azok mennyiségét, nem térünk el a nevezéktani ajánlásoktól. A polimorf dimer és tetramer struktúrák mellett a T26 és a T501 lokuszoknál monomer ismétlődéseket - ún. „A-stretch” - is azonosítottunk, amelyek in/del mutációi mindkét marker esetében hozzájárultak a nagyszámú ún. mikrovariáns allél létrejöttéhez. A T501 marker esetében a számos intermedier (pl.: 9.3, 10.2., 11.2) allél kialakulását a TCTA egységek közé beékelődő TC inszerció is eredményezte. A T26 mikroszatellita 12.1, 13.1 és 15.1 alléljainak monomer A₇ pozíciójában bekövetkező G inszerció a GA dimer ismétlődés kialakulását okozhatta. A

T501 alléljainál korábban leírt monomer stuttererek – amely a POP-4 polimerrel végzett elválasztás során „vállas” csúcsokat eredményez - a hosszú 9-11 nukleotidból álló „A-strech”-nek tulajdonítható. A T26 rövidebb, 6-7 nukleotid hosszú homopolimer A ismétlődése esetében hasonló allélcúcs morfológiát nem figyeltünk meg. A T26 és a T501-hoz hasonló hiperpolimorf komplex lokuszok alléljainak elnevezésére a nemzetközi ajánlások nem adnak szigorú követelményeket (*Gill et al., 1994, 1997a, 2001; Bär et al., 1997; Olaisen et al., 1998, Eichmann et al., 2004*). Az elnevezésnek tükröznie kell az adott allél bázispárban kifejezett hosszát. Mindkét lokusznál a tetramer egységek számát vettük alapul. Az intermedier elnevezések minimalizálása érdekében a monomer ismétlődéseknél egy-egy kiinduló struktúrát vettünk referenciának és csak az attól számított eltéréseket jelöltük az allélnevekben. A T501-es lokusznál a standard homopolimer régió az ortológ fajból (acc.no. *AF442815*) izolált A₁₀ struktúra lett, míg ez a T26-nál az A₇ (*Szabolcsi et al.*).

3.8 A *DeerPlex I-II* markereinek fajspecifikussága

Az ISFG nem-humán mikroszatellita markerekre vonatkozó ajánlása az alkalmazott primerek fajspecifikusságának a vizsgálata (*Linacre et al., 2011*). Ez több szempontból is nagyon fontos. A populációs minták gyűjtése során ugyanis az ember illetve más előlények (pl. gombák, fonálférgek) biológiai anyagmaradványával történő kontamináció nem volt kizárható. Ennek lehetősége genotipizálási nehézségeket okozhat ugyanis a PCR reakciók során a gímszarvasra specifikus primer oligonukleotidok más fajok (akár mikrobák) genomjába is anneálódhatnak és a felszokszorozás során kapott amplikonok csökkenthetik a gímszarvas eredetű STR allélok intenzitását valamint – ritkább esetben – az allélméret tartományokba esve nehezíthetik az allélok típusának meghatározását. A nem gímszarvas genomból felszokszorozott amplikonok és a gímszarvasból származó nem STR eredetű, de allélikus mérettartományba eső amplikonokat mikroszatellita allélként nagyon kis valószínűséggel detektálhattunk, ugyanis a nem allélikus PCR termékek a korábbiakban már említett jellegzetes allél morfológia következtében könnyen kiszűrhetők az analízisből. A populációs minták feldolgozása során kontrollmintákat használtunk (agancsminták esetében annak felületéről biztosított törletminta), amelyekben a nem

STR eredetű amplikonok azonosíthatók, ám ilyeneket egyik kontrollmintában sem detektáltuk. A tervezett gímszarvas specifikus primer szekvenciák más fajokkal esetlegesen fellépő szekvencia egyezését a Genbank adatbázisában (<http://www.ncbi.nlm.nih.gov/genbank/>) ellenőriztük, ám 85 százalék fölötti szekvencia hasonlóságot nem találtunk.

A humán forenzikus gyakorlatban alkalmazott multiplex PCR kitek az ortológ lokuszok szekvencia hasonlósága miatt keresztreakciót adnak más főemlős fajokkal (*Applied Biosystems, AmpFlSTR®NGM™ PCR Amplification Kit User Guide*). A markerek fajspektrumát a *DeerPlex* lokuszokra is le kellett tesztelni. Más, közel rokon patás fajokból ugyanis könnyen amplifikálódhatnak STR allél eredetű PCR termékek. Ennek következtében akár más patás faj részleges vagy teljes genotípusa is a gímszarvas mintákba kerülhetett, megváltoztatva ezzel az allélok eloszlását.

A *DeerPlex I-II* lokuszok fajspecifikusságának vizsgálatát - vagyis a primerek keresztreakcióját – emberi és a leggyakoribb recens patás fajokon is elvégeztük. Tekintetbe véve azt, hogy a szikaszarvas (*Cervus nippon*) és a jávorszarvas (*Alces alces*) Magyarországon igen ritka és az előbbi faj nem is őshonos (*Kászoni, 1999; Boldogh, 2002*), ezeket a szarvas fajokat nem vettük bele a keresztreakció vizsgálatába. Gímszarvas primerekkel PCR termékeket az emberi, a vadsertés és az őz mintákban sem mutattunk ki. A muflon genom esetében három, a szarvasmarha minta esetében egy – nem allélikus mérettartományba eső - csúcsot detektáltunk. A dámszarvas templáton 9 lokusz mutatott keresztreakciót, ellenben a T26 marker nem amplifikálódott. A 10 amplikonból 5 fragmens mérete nem volt beilleszthető egyik allélikus mérettartományba sem, a csúcsok morfológiája azonban tipikusnak volt nevezhető. Allélikus termékeket dámszarvason egyedül a T26 lokusz esetében nem kaptunk. A mikroszatellita markerek tehát a *Cervinae* alcsaládra specifikusak (*Szabolcsi et al.*). A gímszarvas fajspecifikusságot a T26 marker esetében sem lehet kijelenteni, ugyanis a szikaszarvason keresztreakciós vizsgálatokat – a mintavétel nehézségei miatt - nem végeztünk.

Populációgenetikai és statisztikai eredmények

Tekintve azt a két tény, hogy a populáció struktúrája mind a mai napig nem teljesen ismert, valamint azt, hogy a mikroszatellita lokuszok kapcsoltsági csoportjai, a Slate-féle kapcsoltsági térképen lokalizált négy marker kivételével (T108, T156, T172 és a T193), nem ismertek (Slate *et al.*, 2002). Az allélok lokuszon belül és lokuszok között esetlegesen meglévő nem véletlenszerű asszociációját meg kellett vizsgálnunk.

3.9 Hardy-Weinberg egyensúly teszt

A lokuszon belüli véletlenszerű párosodás vizsgálatát a Hardy-Weinberg exakt teszttel végeztük el, aminek szignifikancia szintjét a Bonferroni eljárással korrigáltuk (Szabolcsi *et al.*). A dememorizációs lépések száma az Arlequin szoftvernél 1000, a Genepop szoftver esetében 10000 volt. Az Arlequin lokuszonként 100172 generált tábla adataiból kalkulálta az L_n értéket, a Genepop lokuszonként eltérő számú lépésből (917-23257) becsülte a L_n eloszlást. A lokuszonként kapott p értékeket táblázatos formában közöltük (2. táblázat). A tíz lokusz Hardy-Weinberg egyensúlyát az alpopulációkon (EK, DNY) és az egyesített mintapopuláción (T) teszteltünk, ezért az 5 %-os szignifikancia szintet a Bonferroni összefüggés szerint 0,5 %-ra csökkentettük. A korrigált értéken a tíz lokusznál csak a T501 markernél a teljes és az EK populációban - a Genepop vagy az Arlequin szoftver által kapott p érték alapján - találtunk szignifikáns eltérést a pánmiktikus egyensúlytól. A markereket egyenként vizsgálva, az 5 % szignifikancia szinten a T501 marker alléljainál minden populációs mintában, a C01 markernél az EK és a T108 lokusznál a teljes populációs mintán belül volt elvethető az a hipotézis, hogy az allélok a lokuszon belül random módon párosodnak (2. táblázat). A jelenség összefüggésben állhat a T501 marker hiperpolimorf jellegével. Az allélok számához viszonyított relatíve kis populációk (EK- 48, WP- 100) egyedszám következtében egyes ritka allélokat esetleg nem detektáltunk. A minta populációból „kiesett” allélok a várt és a kalkulált genotípus gyakoriságok közötti különbségen keresztül eltolhatják a Hardy-Weinberg egyensúlyt (NRC, 1996).

2. táblázat A lokuszonként számított Hardy-Weinberg exakt teszt eredménye az EK-észak-keleti, DNY- dél-nyugati és a teljes (T) populációban. *Arlequin*, *Genepop* – a két alkalmazott szoftver. H_O , H_E - megfigyelt és várt heterozigóta gyakoriság, p - a tesztstatisztika értéke.

Lokusz	Populáció EK				Populáció DNY				Egyesített (T) Populáció			
			Arlequin	Genepop			Arlequin	Genepop			Arlequin	Genepop
	H_O	H_E	p	p	H_O	H_E	p	p	H_O	H_E	p	p
T108	0.750	0.780	0.5142	0.5446	0.731	0.810	0.1284	0.0960	0.740	0.822	0.0096	0.0099
T172	0.833	0.879	0.4956	0.4806	0.904	0.870	0.6871	0.6884	0.870	0.891	0.0785	0.0569
T123	0.875	0.856	0.1132	0.1079	0.788	0.827	0.6333	0.6357	0.830	0.859	0.3641	0.3702
C229	0.500	0.555	0.2000	0.1965	0.481	0.447	0.5266	0.5298	0.490	0.500	0.2441	0.2393
T26	0.917	0.921	0.6610	0.6377	0.827	0.881	0.1562	0.2175	0.870	0.914	0.2053	0.1459
C01	0.833	0.868	0.0435	0.0493	0.827	0.840	0.5618	0.5071	0.830	0.875	0.1324	0.1196
T501	0.917	0.940	0.0014	0.0109	0.731	0.886	0.0276	0.0721	0.820	0.920	0.0065	0.0000
T193	0.917	0.904	0.5180	0.5243	0.827	0.900	0.3087	0.3119	0.870	0.914	0.4509	0.4749
T507	0.854	0.882	0.8426	0.8320	0.827	0.821	0.8212	0.8108	0.840	0.868	0.2659	0.3078
T156	0.854	0.867	0.1846	0.1758	0.846	0.910	0.0616	0.0756	0.850	0.903	0.1043	0.1022

3.10 Kapcsoltsági (LE) teszt és marker lokalizáció

Igazságügyi genetikában alkalmazott mikroszatellita markerek kromoszómás lokalizációjának pontos ismerete elengedhetetlen, hiszen a multiplex PCR rendszerek alapkritériuma az alkalmazott markerek független öröklődése (*Butler, 2011*). A teljes gímszarvas genom DNS szekvenciája jelenleg nem ismert (*Genbank*), egyedül a mitokondriális genom szekvenciáját publikálták (*Wada et al., 2010*). A gímszarvas haploid kromoszómaszáma 34 (*Gustavsson and Sundt, 1968*). Slate és munkatársai a termékeny gímszarvas - Dávid szarvas F_1 hibridek visszakeresztezésével kapcsoltsági térképet készített (*Slate et al., 2002*). Több mint 600, I és II típusú markert használtak, amelyeket 33 autoszómás kapcsoltsági csoportra valamint a nemi kromoszómára (PAR - pszeudoautoszómális régió) lehetett elkülöníteni. A szerző az ortológ markerek szinténiájára alapozva gímszarvas – szarvasmarha összehasonlító kapcsoltsági géntérképet is készített, amely alapján a szarvasmarha kromoszómák és a gímszarvas kapcsoltsági csoportok összevethetők. Más szerzők FISH alapú térképezéssel elvégezték a gímszarvas - szikaszarvas – szarvasmarha kromoszóma ortológok

párosítását (*Bonnet et al., 2001*), amely megerősítette Slate és munkatársainak eredményeit. A Slate által készített kapcsoltsági géntérképen a II. típusú markerek között megtalálhatók a *DeerPlex* rendszereinkben alkalmazott T108, T156, T172 és a T193 tetranukleotid mikroszatelliták. Ennek a négy markernek tehát már ismert volt a kromoszómás elhelyezkedése. Az általunk felhasznált további hat polimorf STR lokusz kapcsoltsági viszonyait a szarvasmarha – gímszarvas genom szekvenciájának hasonlóságára alapozott térképezéssel valamint *LD* statisztikai teszttel vizsgáltuk.

A GenBank adatbázisban a teljes szarvasmarha genom DNS szekvenciája annak kromoszómás lokalizációval együtt hozzáférhető. A BLAST programot alkalmaztunk annak érdekében, hogy a 10 tetranukleotid lokusz szarvasmarha ortológját megkeressük. A szarvasmarha kromoszóma gímszarvas párját a Slate-féle összehasonlító térkép felhasználásával azonosítottuk. A BLAST keresés eredményeképpen a 10 STR markerből kilenc ortológját sikerült azonosítani a szarvasmarha genomban. A szekvencia hasonlóságon alapuló keresés során egyedül a T507 lokusz lokalizációja nem volt egyértelműen meghatározható (*Szabolcsi et al.*). A vizsgálat alapján a T26 és a C01 markerek ortológjai a 16-os szarvasmarha kromoszómán lokalizáltak, amely a Slate-féle összehasonlító térkép szerint a gímszarvas 14-es kapcsoltsági csoportjával homológ. A két STR lokusz azonos kromoszómán való elhelyezkedése a vizsgálat alapján nem zárható ki. A T26 marker – ún. „flanking régiójának” - szekvencia azonossága az ortológ párjával azonban relatíve alacsonyabb (79,37%), mint a többi 8 lokusz esetében (százalékos nukleotid egyezés (%): T123:86,24; T501:87,39; T108:88,69; C01:89,9; T172:91,21; T156:93,91; C229:96,54). Ezen kívül, a T26 lokusz kérdéses szarvasmarha ortológja az ismétlődő régió struktúrája miatt egyértelműen nem tekinthető mikroszatellitának. A szarvasmarha ortológ ismétlődő régiójaként a (TATC)₂TGTC(TATC)₂ szekvencia motívumot tudtuk azonosítani. A repeat régió belül a közbeiktatott TGTC egység miatt az ún. „loop” (hurok) képződés és ennek következtében a polimeráz enzim „slipped-strand” mechanizmusa is gátolt lehet, vagyis a T26 szarvasmarha ortológ STR génnek egyértelműen nem definiálható (*Walsh et al., 1996; Clarke et al., 2001; Ellegren et al., 2004*). A szekvencia hasonlósági vizsgálat tehát nem zárja ki annak lehetőségét, hogy a T26 lokuszhoz igen hasonló szarvasmarha szekvencia nem az ortológ mikroszatellita gén.

A 10 marker egymástól független öröklődésének vizsgálatát *LD* statisztikai teszttel is elvégeztük. A teszt során összesen 45 lokuszpár kapcsoltságát vizsgáltuk a három - EK, DNY és ezek együtteséből (T) álló – populációs mintán. Az EK populációs mintában öt, a DNY populációs mintában két lokuszpár, az összesített populációs mintában nyolc lokuszpár esetében volt elvethető a kapcsoltsági egyensúly az 5 % szignifikancia szinten. Olyan lokuszpárt, amely mindhárom populációban a kapcsoltsági egyensúlytalanság állapotában volt, nem mutattunk ki (*Szabolcsi et al.*). Ez ugyancsak cáfolja azt az eredményt, hogy a T26 - C01 markerek egy kapcsoltsági csoportba tartoznak. A teljes minta populációban az 5 % szignifikancia szinten nyolc lokuszpár tért el a kapcsoltsági egyensúlytól. A Bonferroni korrekciót alkalmazva ($p < 0,0011$) csak a teljes mintapopulációban mutattunk ki három olyan lokuszpárt, amelyek kapcsoltsági egyensúlytalanságban voltak (3. táblázat). Két lokusz között *LD* állapot azonban nemcsak fizikai kötődés, hanem a mintapopuláció tagoltsága következtében is kialakult HW egyensúlytalanságból is adódhat (*Hartl and Clarck, 2007*).

A markereket izoláló szerzők nem publikálták sem a donor állatok számát, sem azok nemét. A mikroszatelliták polimorfizmusát 28 – 43 referencia egyeden vizsgálták, azonban kapcsoltsági vizsgálatokat nem végeztek. A referencia állatok neméről sem állt rendelkezésre adat (*Jones et al., 2002; Meredith et al., 2004*). A markerek esetlegesen az X vagy Y kromoszómán való lokalizálódását vizsgálnunk kellett. A 10 marker újra szekvenálását ezért szarvastehénből amplifikált lokuszokon végeztük annak érdekében, hogy az Y kromoszómás STR markereket már ennél a lépésnél kizárjuk. A populációs polimorfizmus vizsgálatok során nem találtunk egyetlen olyan markert sem, amely hím egyedeknél csak homozigóta genotípusú lett volna.

A vizsgálatok során minden marker lokalizációjánál kizárható volt az X kromoszóma és Y kromoszóma Y specifikus régiója. Azonban Y kromoszómán elhelyezkedő STR marker esetében szarvasbikán és szarvastehéneken is detektálhatunk heterozigóta genotípust, amennyiben a marker az Y kromoszóma X kromoszómával homológ, rekombinálódásra képes pszeudoautoszómális szakaszán lokalizált (PAR régió). A nemzetközi szakirodalomban publikált adatok szerint az emlősök PAR régiója relative kisméretű a teljes Y kromoszóma méretéhez képest (5%, *Butler, 2011*). Wan-Sheng és munkatársai által vizsgált 57 szarvasmarha Y kromoszómás mikroszatellita közül mindössze 6 STR bizonyult PAR régió specifikusnak (*Liu et al., 2002*). Az utóbb

említett publikált adatokból következően, PAR Y-STR lokusz nagyon kis valószínűséggel kerülhetett bele a 10 markert tartalmazó készletünkbe, azonban ennek lehetősége nem zárható ki.

Összegezve a vizsgálatok eredményét, a lokuszokat autoszómás öröklődésűnek tekintjük és 9 STR lokusztól azt feltételezzük, hogy a gímszarvas genomban a nemi kromoszómákon kívül, külön kapcsoltsági csoportba tartoznak. A T26-C01 STR lokusznál a genetikai kapcsoltságot nem zárhatjuk ki, ezért az igazságügyi genetikai számításokhoz a C01 lokusz kihagyását ajánljuk.

STR	T108			T172			T123			C229			T26			C01			T501			T193			T507			T156		
	E	D	T	E	D	T	E	D	T	E	D	T	E	D	T	E	D	T	E	D	T	E	D	T	E	D	T	E	D	T
T108				-	-	-	-	-	-	-	-	-	-	-	-	-				-	-	-	-	-	-	-	-	-	-	-
T172	-	-	-				-	-	-	-	-	-	-	-	-				-	-	-	-	-	-		-		-	-	-
T123	-	-	-	-	-	-				-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		-	-	-	-
C229	-	-	-	-	-	-	-	-	-				-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T26	-	-	-	-	-	-	-	-	-	-	-	-				-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C01	-	-		-		-	-	-	-	-	-	-	-	-	-				-				-		-	-		-	-	-
T501	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-							-	-	-	-		-	-	-
T193	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				-	-	-	-				-			-	-
T507	-	-	-		-		-	-		-	-	-	-	-	-	-	-		-	-			-					-	-	-
T156	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		-	-		-	-			

3. táblázat Lokuszpárok között szignifikánsan (: $\alpha = 0,05$; : Bonferroni korrekció esetén $\alpha = 0,0011$) elvethető kapcsoltsági egyensúlyok (*LD*). A kapcsoltsági egyensúlyban (*LE*) lévő lokuszpárok jelölése: – , E - észak-keleti, D - dél-nyugati és a T - teljes populáció.

3.11 A 10 autoszómás STR lokusz genetikai diverzitása

A nagyfokú genetikai diverzitást mutató markerek nagymértékben növelik mikroszatellita készletünk igazságügyi egyedazonosításhoz szükséges individualizáló erejét. A genetikai diverzitás mérésének sokféle módszer létezik, mi a két legáltalánosabbat, a Nei-féle géndiverzitást (h vagy H_E) és a polimorfizmus információ tartalmat (PIC) választottuk. A Nei-féle érték becslése megegyezik a mintában előforduló allélok számával korrigált heterozigóta gyakoriság kalkulálásával. A PIC számítása a korrigálatlan heterozigóta gyakorisághoz nagyon hasonló, azonban a H értékből levonja két azonos heterozigóta genotípus párosodási valószínűségét. Könnyen belátható, hogy a PIC értéke mindig kisebb, mint a heterozigóták gyakorisága vagy a Nei-féle géndiverzitás (4. táblázat). Egy lokusznak minél több allélja van, a PIC és a H értéke annál jobban közelít egymáshoz (Botstein et al., 1980; Ott, 1999). Ez jól megfigyelhető a két legmagasabb polimorfizmus fokkal rendelkező T26 és a T501 lokusznál. A magas allélszám (21 és 23) következtében a PIC és a H értéke kevesebb, mint 1,5 %-ban tért el egymástól (4. táblázat). A h és a PIC érték alapján mindhárom mintapopulációban a C229 genetikai diverzitása a legalacsonyabb ($H_E = 0,556$ (EK); $0,447$ (DNY); $0,5$ (T); $PIC = 0,501$ (EK); $0,404$ (DNY); $0,453$ (T)), a T501 markernél pedig a legmagasabb ($H_E = 0,94$ (EK); $0,886$ (DNY); $0,92$ (T); $PIC = 0,926$ (EK); $0,866$ (DNY); $0,909$ (T)). A C229 markernél becsült alacsony h és PIC érték nemcsak a relative kevés allélnak (5), hanem a 15-ös allél rendkívül magas populációs gyakoriságának ($0,625$ (EK); $0,721$ (DNY); $0,675$ (T)) együttesen tulajdonítható. Ezek az eredmények azt sugallták, hogy mindhárom mintapopulációban – vagyis a magyarországi gímszarvas állományban - a 10 STR marker közül a T501 lokusz bír a legnagyobb igazságügyi megkülönböztető erővel (PD), amelynek értéke mindhárom populációban valóban magas volt ($PD > 0,95$). A legmagasabb PD értéket azonban az teljes és a délnyugati populációs mintában a T193 markernél ($0,9652$; $0,9776$) az északkeleti populációban pedig a T26 markernél kalkuláltuk ($0,9688$).

	<i>N</i>			<i>H₀</i>			<i>H_E</i>			<i>PIC</i>			<i>PD</i>		
	<i>E</i>	<i>D</i>	<i>T</i>	<i>E</i>	<i>D</i>	<i>T</i>	<i>E</i>	<i>D</i>	<i>T</i>	<i>E</i>	<i>D</i>	<i>T</i>	<i>E</i>	<i>D</i>	<i>T</i>
T108	9	8	10	0,750	0,731	0,740	0,781	0,811	0,822	0,741	0,776	0,794	0,9080	0,9216	0,9364
T172	12	12	13	0,833	0,904	0,870	0,879	0,870	0,891	0,857	0,848	0,876	0,9575	0,9534	0,9678
T123	10	9	10	0,875	0,789	0,830	0,856	0,827	0,859	0,829	0,798	0,838	0,9288	0,9401	0,9560
C229	5	4	5	0,500	0,481	0,490	0,556	0,447	0,500	0,501	0,404	0,453	0,7352	0,6583	0,6982
T26	18	17	21	0,917	0,821	0,870	0,922	0,882	0,914	0,905	0,862	0,903	0,9688	0,9578	0,9762
C01	10	12	12	0,833	0,821	0,830	0,868	0,840	0,875	0,844	0,813	0,857	0,9436	0,9423	0,9632
T501	22	19	23	0,917	0,712	0,810	0,940	0,885	0,920	0,926	0,866	0,909	0,9644	0,9593	0,9748
T193	14	14	15	0,917	0,821	0,870	0,904	0,900	0,914	0,885	0,881	0,902	0,9635	0,9652	0,9776
T507	11	9	11	0,854	0,821	0,840	0,882	0,821	0,868	0,860	0,791	0,851	0,9592	0,9357	0,9606
T156	12	15	15	0,854	0,846	0,850	0,867	0,910	0,903	0,843	0,893	0,889	0,9444	0,9645	0,9710

4. táblázat A lokuszonként kimutatott allélszám (*n*), a megfigyelt és várt heterozigóta gyakoriság (*H₀*, *H_E*), a polimorfizmus információ tartalom (*PIC*), a diszkrimináló erő (*PD*) az E- észak-keleti, D- dél-nyugati és a T- teljes populációban.

A minimális allélfrekvencia értékei az EK populációban 4,7 és 6,3 % között, a délnyugati populációban 4,1 és 5,3 % tartományban mozogtak (5. táblázat). Az északkeleti populáció viszonylag magas, 6,3 % minimális allélgyakorisági értéke az alacsonyabb mintaszámnak (48) volt köszönhető.

A két mintapopulációt összehasonlítva a H , PIC , PD értékeinek maximumát az északkeleti populációban mutattuk ki, vagyis a két populáció közül az utóbb említett tekinthető a genetikailag változékonyabbnak és egyben az igazságügyi egyedazonosítás hatékonyságának szempontjából a kedvezőbbnek is (Szabolcsi *et al.*). Ennek részleteit egy későbbi fejezetben tárgyaljuk.

5. táblázat A populációs mintákban vizsgált, 10 lokuszra vonatkozó megfigyelt és várt heterozigóta gyakoriság (H_0 , H_E), polimorfizmus információ tartalom (PIC), diszkrimináló erő (PD) és minimális allélgyakoriság (p_{min}) szélső értékei.

	H_0	H_E	PIC	PD	P_{MIN}
MAXIMUM	0,917	0,94	0,926	0,9688	0,0626
LOKUSZ	T26,T501,T193	T501	T501	T26	T26,T501,T193
POPULÁCIÓ	ÉSZAKKELET	ÉSZAKKELET	ÉSZAKKELET	ÉSZAKKELET	ÉSZAKKELET
MINIMUM	0,48	0,447	0,404	0,6583	0,0412
LOKUSZ	C229	C229	C229	C229	C229
POPULÁCIÓ	DÉLNYUGAT	DÉLNYUGAT	DÉLNYUGAT	DÉLNYUGAT	DÉLNYUGAT

A vizsgálati eredmények alapján felvetődött annak a lehetősége is, hogy a két populáció között szignifikáns allélfrekvencia különbségeket mutathassunk ki, vagyis a populációk genetikailag differenciáltak lehetnek.

Minden mikroszatellita lokusz alléljainak populációs eloszlását táblázatos formában közöltük (F.4-F.13. ábrák). Vizuálisan is megállapítható, hogy különösen a T108 és a C01 lokuszokon mért allélgyakoriságok és allélok eloszlása tekintetében szignifikáns különbségek vannak a két populáció között. A C229 alléljainak eloszlásában viszont nincs szemmel látható különbség.

Ha a populáció divíziókra bontható, vagyis genetikailag differenciált, akkor a h becslése nem közelít a HWE egyensúlyi gyakorisághoz, ugyanis a géndiverzitás ebben az esetben az alpopulációk (divíziók) átlaga (Nei, 1973). Ha a populáció beltenyésztett, akkor a populációban a párválasztás nem véletlenszerű és ebben az esetben a h alulbecsli a tényleges heterozigóták gyakoriságát (DeGiorgio et al., 2010).

3.12 Genetikai differenciálódás és beltenyészet

A gímszarvas populáció genetikai struktúrájának vizsgálatához F-statisztikát végeztünk. A populáció genetikai struktúrájának vizsgálata igazságügyi aspektusból is nagyon fontos, ugyanis a több lokuszos genotípus gyakoriság számításának alapja a Hardy-Weinberg törvény. A mintapopulációkban az allél frekvenciák megfigyelt és várt értékei között azonban szignifikáns eltérés is lehetséges, ami a több lokuszos genotípus gyakoriság elfogult becsléséhez vezethet. Az eltérés mértékét meg kellett határozzuk és a korrekciós paramétereket fel kellett használjuk az Ayres és Overall által kidolgozott gyakorisági formulákba (Ayres and Overall, 1999).

Ha a populáció (mint metapopuláció) alpopulációkra (ún. démekre) bontható abban az esetben, a populációban a heterozigóták aránya az alpopulációk átlaga, amely sosem nagyobb, mint a Hard-Weinberg egyensúly (HWE) állapotban (Butler, 2011). A HWE tehát nem feltétel, ugyanis két HWE állapotban lévő populáció egyesítése is mérhető heterozigóta redukciót illetve homozigóta többletet okozhat. Heterozigóta többletet más mikroevolúciós hatások okozhatnak, mint például a szelekció, a mutáció, negatív asszortatív párosodás, a hibrid vigor és a kis populációban a két nem között kialakuló allélfrekvencia különbségek (Hedrick, 2011). A heterozigóta arány csökkenésének a mértékét a Nei-féle formulával is számszerűsíthetjük (Nei, 1973):

$F_{ST} = \frac{H_T - \bar{H}_S}{H_T}$ ami hozzávetőlegesen megegyezik a varianciaanalízissel kalkulált

$F_{ST} = \frac{\sigma_a^2}{\sigma_T^2}$ értékkel (Excoffier et al., 1992; Hartl and Clark, 2007). H_T - a pánmiktikus

teljes populációban várt heterozigóták gyakorisága, \bar{H}_S - a heterozigóták várt gyakoriságának átlaga a pánmiktikus alpopulációkban.

A beltenyészet hatása a genotípus eloszlásra igen hasonló ahhoz, mint amit az alpopulációk egyesítése következtében kialakult inhomogenitás okoz. A populáción belül kialakult nem véletlenszerű párosodás – a rokon egyedek valószínűbb párválasztása - következtében a pánmiktikus egyensúlyhoz képest homozigóta többlet és ezzel együtt heterozigóta hiány keletkezik. Ennek mértékét a Nei-féle formulával is számszerűsíthetjük (Nei, 1973):

$$F_{IS} = \frac{\bar{H}_s - H_I}{\bar{H}_s} \text{ amit variancia analízissel a } F_{IS} = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_c^2} \text{ formulával fejezhetünk ki.}$$

H_I - a heterozigóták megfigyelt aránya az alpopulációban.

A két mikroevolúciós hatás együttesének befolyását a populáció egyensúlyi heterozigóta gyakoriságára a következőképpen fejezhetjük ki:

$$(1 - F_{IS})(1 - F_{ST}) = (1 - F_{IT}) \text{ ahol, } F_{IT} = \frac{H_T - H_I}{H_T} \text{ ami variancia analízissel (Excoffier et}$$

al., 1992) kifejezve $F_{IT} = \frac{\sigma_a^2 + \sigma_b^2}{\sigma_T^2}$. Az F_{IS} , F_{ST} és az F_{IT} értékeket AMOVA-t

alkalmazva minden lokuszra külön kiszámítottuk, és azokat a hozzájuk tartozó p értékkel együtt táblázatos formában közöltük (6. táblázat).

Az AMOVA elemzés során lokuszonként kapott F_{ST} értékeket az EK-DNY populációpárra kalkuláltuk. A C229 lokusznál a két vizsgált populáció alléleloszlása között nem volt szignifikáns különbség. A többi kilenc marker esetében a populációk között szignifikáns allélvariancia különbségeket mutattunk ki. Hasonlóan a többi markerhez, a T501 lokusz esetében ez a genetikai elkülönülés is szignifikáns volt, de elhanyagolható mértékű (hozzávetőlegesen 1,5 %). A genetikai differenciálódás mértéke a T108 lokusz esetében mutatta a legnagyobb mértéket ($F_{ST} = 0,0611$), vagyis a teljes allélvariancia több mint 6 %-ban a populációk közötti különbségből adódott.

A beltenyészet becslése egyedül a T501 mikroszatellitánál a DNY és a teljes populációban volt szignifikáns és igen jelentős, a többi markernél a véletlen szerepe túlságosan jelentős volt ahhoz, hogy a heterozigóta genotípusok pánmiktikus egyensúlyi állapottól tapasztalt populáción belüli eltérésének mértékét lokuszonként vegyük figyelembe. A T501 STR markernél az átlagos populáción belüli allélvariancia több mint 10 %-a volt a teljes allélvarianciának, ami igen jelentős hányad. A lokuszonként

becsült nem szignifikáns F_{IS} értékek eredményezték azt, hogy az F_{IT} paraméterek esetében a nullhipotézis ($H_0 : F_{IT} = 0$) nem volt elvethető az 5 %-os konfidencia szinten. A lokuszonként becsült populációs paraméterek eredményeit összegezve kijelenthető hogy a vizsgált gímszarvas populációkban becsült populációs paraméterek közül a profilgyakorisági értékek kalkulációjához az F_{ST} paraméter lokuszonkénti értéke is figyelembe vehető. A genetikai tagoltság mértéke a T108, C01, T123 és a T507 STR markereknél volt a legnagyobb ($> 4 \%$). A vizsgálatok alátámasztják a C01 és T108 lokuszok mellékletben közölt (F.9., F.13. ábrák), első pillantásra is megállapítható populációs alléleloszlás különbségét.

A populációk között és azokon belül lezajlott mikroevolúciós hatásokat leginkább akkor tudjuk becsülni, ha mind a 10 lokusz adatait figyelembe vesszük. Ennek kivitelezéséhez ún. „globális” molekuláris varianciaanalízist (AMOVA) végeztünk. A kalkulált varianciahányadosokat és a permutációs teszteket ebben az esetben 10 lokuszos genotípus/haplotípus adatokkal végzi a szoftver. A kapott eredményeket a 7. táblázatban közöltük.

7. táblázat A „globális”, 10 mikroszatellita lokuszra számított AMOVA eredménye a

két gímszarvas populációban. $n = \frac{2N - \sum_{p \in P} \frac{2N_p^2}{N}}{P - 1}$, ahol P az alpopulációk száma, N a teljes egyedszám és N_p a p -dik alpopulációt alkotó egyedek száma.

GENETIKAI/ALLÉL VARIANCIA FORRÁSA	NÉGYZETÖSSZEG (10 LOKUSZ)	VARIANCIA KOMPONENS	VARIANCIA SZÁZALÉK
POPULÁCIÓK KÖZÖTT	18.924	0.14632 ($n\sigma_a^2 + 2\sigma_b^2 + \sigma_c^2$)	3.39759
POPULÁCIÓN BELÜL	422.901	0.15516 ($2\sigma_b^2 + \sigma_c^2$)	3.60295
EGYEDEK KÖZÖTT	400.500	4.00500 (σ_c^2)	92.99945
TELJES VARIANCIA	842.325	4.30648 (σ_T^2)	100.00000

LOKUSZ	N	F_{IS}	P	F_{IS}	P	F_{IS}	P	F_{ST}	P	F_{IT}	P
		EK		DNY		T					
T108	10	0.0395	0.37243	0.0992	0.08798	0.0711	0.07820	0.0611	0.00000	0.1279	0.01857
T172	13	0.0519	0.20723	-0.0395	0.84066	0.0046	0.50635	0.0366	0.00000	0.0411	0.27077
T123	10	-0.0228	0.69892	0.0472	0.27273	0.0130	0.41251	0.0408	0.00000	0.0533	0.19941
C229	5	0.1008	0.19550	-0.0760	0.82405	0.0185	0.44966	0.0027	0.44184	0.0211	0.42522
T26	21	0.0048	0.54839	0.0620	0.14761	0.0339	0.17302	0.0296	0.00000	0.0625	0.07722
C01	12	0.0403	0.31378	0.0162	0.44575	0.0279	0.27077	0.0478	0.00000	0.0744	0.10068
T501	23	0.0255	0.30694	0.1767	0.00098	0.1019	0.00293	0.0155	0.04594	0.1158	0.00000
T193	15	-0.0147	0.67840	0.0817	0.07625	0.0353	0.16911	0.0260	0.00098	0.0604	0.08309
T507	11	0.0319	0.33431	-0.0074	0.62952	0.0122	0.41447	0.0404	0.00000	0.0520	0.22385
T156	15	0.0148	0.37243	0.0708	0.08798	0.0446	0.13196	0.0284	0.00196	0.0718	0.06354

6. táblázat A lokuszonként számított AMOVA eredménye az EK- észak-keleti, DNY- dél-nyugati és az egyesített (T) populációban. n – a megfigyelt allélszám, p – a tesztstatisztika értéke.

A 100 egyedből álló gímszarvas mintapopulációban a varianciaanalízissel becsült átlagos beltenyészet paraméterre (F_{IS}) 0,0373, átlagos szubdivízió paraméterre (F_{ST}) 0,03398, az átlagos teljes beltenyészet paraméterre (F_{IT}) 0,07001 értéket kaptunk (Szabolcsi *et al.*). A HW egyensúlyi állapottól a heterozigóta gyakoriság átlagosan 7 %-os csökkenést mutatott. A szoftver által elvégzett 1023 permutáció során a számított F_{ST} és F_{IT} értékek egyik esetben sem haladták meg a kapott értékeket (vagyis $p = 0,0000$). Az F_{IS} paraméter a permutációs teszt során az 1023 eset 0,293 százalékában haladta meg vagy érte el a 0,0373 értéket. A nullhipotézist elvethetjük, vagyis az 5%-os konfidencia szinten tehát minden paraméter szignifikánsnak tekinthető. Tekintetbe véve, hogy a lokuszonként becsült F_{IS} értékek nem voltak szignifikánsak, a profilgyakoriság korrekciójánál a 10 lokuszos globális AMOVA során kapott értékeket vesszük figyelembe. A szubdivízió paraméter esetében a lokuszonként kimutatott szignifikáns értékek is számításba vehetők.

Más európai gímszarvas populációk között más szerzők is kimutattak genetikai különbségeket. Kuehn és mtsai kilenc bajorországi, öt svájci, egy türingiai, egy csehországi (Boehmerwald), egy olaszországi (Vinschgau) és egy lichtensteini gímszarvas populáció genetikai különbségét vizsgálták 19 majd 18 mikroszatellita allél polimorfizmusának felmérésével (Kuehn *et al.*, 2003, 2004). Az eredmények azt mutatták, hogy a 17-19. századig lezajlott palacknyak effektus és az alapító hatás következtében bizonyos populáció párok között magas genetikai különbség mérhető. A legmagasabb átlagos allélfrekvencia különbséget az észak és dél-bajorországi populáció között mutatták ki, amely majdnem 20 % volt ($F_{ST} = 0,1910$; $p < 0,0009$). A jelen tanulmányban vizsgált két magyarországi populáció közötti genetikai tagoltság mértéke a - Kuehn és mtsai által vizsgált - nyugat-európai populációkhoz képest relative kismértékű ($F_{ST} = 0,03398$; $p = 0,0000$).

3.13 A *DeerPlex I* és *II* individualizáló ereje

A 10 mikroszatellitára becsült átlagos egyezési valószínűség (*probability of match, pM*) vagy az átlagos azonosság valószínűségnek (*probability of identity, PI_{ave}*) a marker készletünk individualizáló erejét számszerűsíti, vagyis azt, hogy a 10 STR markert tartalmazó *DeerPlex I* és *II* milyen mértékben alkalmas a magyarországi gímszarvas

egyedek genetikai úton történő elkülönítésére és azonosítására. Ennek becsléséhez figyelembe kell vennünk azt a tényt, hogy a populáció struktúrált, és az egyedek a populációkon belül bizonyos mértékig rokonai egymásnak. Ha a megfigyelt - teljes populációban mért - allélgyakoriságokból több lokuszos genotípus gyakoriságokat multiplikálunk, akkor pánmiktikus egyensúlytól való eltérés mértéke átlagosan 7% lesz ($F_{IT} \approx 0.07$), vagyis a HWE állapotból kalkulált individualizáló erőt (várt érték) a teljes populációban becsült inhomogenitás és a beltenyészet átlagosan 7%-ban csökkenti. A valós pM/PI_{ave} értékek kalkulációját a korrekciós paraméterek alkalmazásával (F_{IS} , F_{ST}) az API-CALC v1.0 szoftverrel végeztük el (Ayres and Overall, 2004). A lokuszonként kapott értékeket a 8. táblázatban foglaltuk össze.

8. táblázat A lokuszonként számított átlagos egyezési/azonossági valószínűségek (pM/PI_{ave}) az EK- észak-keleti, DNY- dél-nyugati és a teljes (T) populációban.

LOKUSZ	PI_0			PI_E		
	EK	DNY	T	EK	DNY	T
T108	0.0727	0.0603	0.0541	0.0828	0.0649	0.0725
T172	0.0222	0.0279	0.0224	0.0296	0.0322	0.0357
T123	0.0514	0.0415	0.0343	0.0405	0.0531	0.0517
C229	0.2491	0.3288	0.2947	0.2515	0.3538	0.3257
T26	0.0106	0.0234	0.0139	0.0142	0.0274	0.0259
C01	0.0363	0.0392	0.0271	0.0343	0.0469	0.0433
T501	0.0151	0.0234	0.0158	0.0090	0.0265	0.0239
T193	0.0160	0.0158	0.0125	0.0200	0.0215	0.0262
T507	0.0204	0.0460	0.0297	0.0290	0.0564	0.0459
T156	0.0355	0.0166	0.0192	0.0349	0.0176	0.0308

Annak a valószínűsége, hogy a teljes populációban véletlenszerűen kiválasztott két egyedben ugyanazt a 10 lokuszos genotípust figyelhetjük meg $2,6236 \times 10^{-15}$ ($F_{IS} = 0,0373$; $F_{ST} = 0,0340$). Az átlagos egyezési/azonosság kombinált valószínűség az EK ($F_{IS} = 0,0241$) és a DNY ($F_{IS} = 0,0498$) populációkban $2,2194 \times 10^{-15}$ és $2,8551 \times 10^{-14}$ (Szabolcsi et al.). A 3.10 fejezetben tárgyalt függetlenségi vizsgálatok eredménye alapján a T26-C01 markerek fizikai kapcsoltsága nem zárható ki, ezért a pM/PI_{ave} értékeket, a C01 mikroszatellita kihagyásával 9 függetlennek tekintett markerre is

kiszámítottuk (ld 9. táblázat). A T26 mikroszatellita lokusz genotípusa a fajspecifikussági vizsgálatok alapján rendkívül fontos, valamint az átlagos pM/PI_{ave} értékei minden populációban alacsonyabbak, mint a C01 markernél.

9. táblázat A C01 STR marker kihagyásával számított kombinált, várt átlagos egyezési/azonossági valószínűségek (pM/PI_{ave}) az EK- északkeleti, DNY- a délnyugati és a T- teljes populációban.

9 STR MARKERRE KOMBINÁLT PI_{AVE}		
EK	DNY	T
6,4717 x 10 ⁻¹⁴	5,8495 x 10 ⁻¹³	1,0056 x 10 ⁻¹²

Mind a 10 mind a 9 lokuszos kombinált átlagos egyezési/azonossági valószínűség értékek az észak-keleti populációban bizonyultak a legalacsonyabbnak, ami a 3.11 fejezetben tárgyalt populáción belüli, magasabb fokú genetikai diverzitásnak köszönhető. Minél magasabb a géndiverzitás annál alacsonyabb az egyezési valószínűség. Mindhárom mintapopulációban a C229 pM/PI_{ave} értéke bizonyult a legmagasabbnak. A lokuszonként számolt pM/PI_{ave} értékekből jól kivehető, hogy azok nemcsak a genetikai diverzitás paramétereitől (h , PIC) függnek, hanem a lokuszok populáción belüli alléleloszlásától is. A géndiverzitás ugyanis nem kalkulál a heterozigóták eloszlásával. A DNY populációs mintában a 14 alléllal és 0,827 H_0 rendelkező T193 marker alacsonyabb megfigyelt pM/PI_{ave} valószínűségi értéket mutattunk ki, mint a 15 alléllal és 0,846 H_0 rendelkező T156 mikroszatellita marker esetében. Az EK, DNY és a teljes (T) populációban a kombinált átlagos igazságügyi megkülönböztető erő (PD_{komb}) igen magasnak bizonyult, nagyobb, mint 0,99999999. Tehát annak az esélye, hogy bármely populációban két gímszarvas egyedet véletlenszerűen kiválasztva - a *DeerPlex I* és *II* markerekkel vagy annak kilenc lokuszra csökkentett készletével - egymástól genetikai vizsgálattal el lehet különíteni, 99,999999 %.

A két mintapopulációt vizsgálva, a kombinált átlagos pM/PI_{ave} és PD paraméterei azt sugallják, hogy a relatíve magasabb géndiverzitás következtében az igazságügyi

egyedazonosítás szempontjából az EK alpopuláció kedvezőbb, ugyanis az északkeleti populáció egyedeit vizsgálva átlagosan alacsonyabb pM/PI_{ave} és magasabb PD értékeket kaphatunk.

A marker készlet individualizáló erejét nemcsak az átlagos pM/PI_{ave} értékekkel fejezhetjük ki, hanem a teoretikusan „leggyakoribb” genetikai profil egyediségével (un. „uniqueness”). Természetesen egyetlen genetikai profil sem tekinthető teljesen egyedinek, ugyanis a genetikai profil az egyeden belül fellelhető teljes genetikai diverzitás egy igen kicsiny része. Azonban a genetikai profil előfordulási gyakorisága olyan kicsiny lehet, hogy az egy adott populáción belül csak egyetlen egyedből mutatható ki. A teoretikusan „leggyakoribb” célprofil az egyed azonosítása szempontjából a legkedvezőtlenebb eset ugyanis ennek az egyedisége a legalacsonyabb, vagyis ez a genotípus az, amely a legnagyobb valószínűséggel fordulhat elő kétszer a populációban. Amennyiben ennek a profilnak a becsült gyakorisága és konfidencia intervalluma nagyságrendekkel kisebb a teljes gímszarvas állomány (N) méretének reciprokához képest vagy másképpen fogalmazva, ha a „leggyakoribb” genetikai profil LR értéke nagyságrendekkel nagyobb a világon létező összes gímszarvasnál, akkor nagy bizonyossággal kijelenthető, hogy minden *DeerPlex I-II* profil egyedi. A „leggyakoribb” genetikai profil becsléséhez a lokuszok független öröklődését feltételezzük és figyelembe vesszük a HWE állapottól becsült eltérés mértékét (F_{ST} , F_{IS}) valamint a minta egyedszámát is (n). A populációnként becsült teoretikusan „leggyakoribb” genetikai profilt és annak 95%-os konfidencia intervallumát a 10. táblázatban közöljük.

Milyen mértékű legyen a különbség a genetikai profil LR értéke és a teljes gímszarvas állomány között annak eldöntésére, hogy az adott profil egyedinek tekinthető-e, nem lehet egyértelműen determinálni. Ennek csak a mértékét lehet számszerűen kifejezni. Egy adott célprofil esetében, ha a populációt N egymással nem rokon egyed alkotja, akkor annak a valószínűsége, hogy az N egyedből legalább egy rendelkezzen a célprofillal NP_{cel} , ahol P_{cel} a célprofil populáción belül kalkulált gyakorisága. Tehát annak a valószínűsége, hogy az adott célprofil egyedi, legalább $1 - NP_{cel}$. Ha nem egy adott célprofil, hanem a teoretikusan „leggyakoribb” profil egyediségével számolunk, akkor a marker készletünk adott populáción belül értelmezett minimális egyediségét kapjuk meg, ami $1 - NP_{max}$.

10. táblázat A populációnként számított teoretikusan „leggyakoribb” és „legritkább” 9 és 10 lokuszos *DeerPlex I-II* genetikai profil az EK- észak-keleti, DNY- dél-nyugati és a teljes populációban. P_{max} , P_{min} – a teoretikusan „leggyakoribb” és „legritkább” genetika profil gyakorisága (pM) és *k.i.* - azok 95 %-os konfidencia intervallumának két szélső értéke. $P_{C01\ max}$ – a C01 kihagyásával 9 lokuszra kalkulált maximális profilgyakoriság.

	<i>EK</i>		<i>DNY</i>		<i>T</i>	
C229	15	15	15	15	15	15
T26	15.1	12.1	17	13	13	13
T193	20	13	13	11	20	13
T108	7	8	12	9	7	7
T123	19	16	13	13	19	13
T501	10.2	10.1	11.2	6.1	11.2	11.2
T172	15	5	14	14	15	14
T156	6	14	5	12	6	14
C01	27	18	20	22	20	20
T507	5	8	8	8	8	8
P_{max}	$1,2548 \times 10^{-11}$		$1,2728 \times 10^{-10}$		$3,3654 \times 10^{-11}$	
<i>k.i.</i>	$6,7333 \times 10^{-11}$		$5,9947 \times 10^{-10}$		$8,5903 \times 10^{-11}$	
	$2,3384 \times 10^{-12}$		$2,7023 \times 10^{-11}$		$1,3185 \times 10^{-11}$	
$P_{C01\ max}$	$1,6459 \times 10^{-10}$		$9,9260 \times 10^{-10}$		$3,9661 \times 10^{-10}$	
P_{min}	$5,1241 \times 10^{-24}$		$5,2475 \times 10^{-24}$		$6,9319 \times 10^{-21}$	
<i>k.i.</i>	$1,9927 \times 10^{-22}$		$7,9736 \times 10^{-23}$		$1,5577 \times 10^{-20}$	
	$1,3176 \times 10^{-25}$		$3,4535 \times 10^{-25}$		$3,0847 \times 10^{-21}$	

A „leggyakoribb” profil számításakor figyelembe vettük az adott populáción belül létező beltenyészet (F_{IS}) és az esetlegesen kimutatható inhomogenitás (F_{ST}) mértékét is. Az eltérő alléleloszlásnak, valamint a populációs paraméterek (F_{IS} , F_{ST} , F_{IT}) genotípus gyakoriságot módosító hatása következtében minden populációban különböző 10 lokuszos genetikai profilt kaptunk. A maximális profil gyakoriságok konfidencia intervallumainak két szélső értéke minden populációban egy nagyságrendnyi különbséget mutatott, ami szűk tartománynak számít (*Chakraborty et al., 1993*). Az N becsléséhez több vizsgálati eredményt is figyelembe kellett veyünk. Az egyediség számításához igen fontos tényező, hogy a marker készlet nemcsak a *Cervus elaphus hippelaphus* alfajon működik. A 3.8 fejezetben tárgyalt fajspecifikussági vizsgálatok alapján, a T26 markert sikeresen felsokszorozó *DeerPlex I-II* és ennek 9 lokuszos

részleges genotípusának minimális egyediségéhez nem kell számításba vegyük a dámszarvas európai létszámát. Az N becsléséhez azonban figyelembe kell vennünk a keresztreakcióra nem tesztelt *Cervus nippon* (szikaszarvas) egyedeit is, amelyen a teljes *DeerPlex I-II* vagy 9 lokuszos profilt adhat. Nem kalkulálhatunk kizárólag a magyarországi vagy az európai gímszarvas állománnyal. A 2010-ben publikált adatok alapján (*Burbaité and Csányi, 2010*) az európai gímszarvas populáció 1,7 millióra tehető, azonban a korábban említett okok miatt számításba kell vennünk a szikaszarvas állományát is, ezért 2 millió szarvas egyeddel (N) számoltunk.

A *DeerPlex I-II* marker készlet és a 9 lokuszra redukált multiplex rendszer minimális egyediségét a 11. táblázatban közöltük.

11. táblázat A populációnként számított *DeerPlex I-II* és C01 lokusz kihagyásával alkotott 9 markert tartalmazó multiplex rendszerek becsült minimális egyedisége az EK- észak-keleti, DNY- dél-nyugati és a teljes (T) populációban.

	<i>EK</i>	<i>DNY</i>	<i>T</i>
<i>I-NP_{MAX}</i>	$1 - 2,51 \times 10^{-5}$	$1 - 2,55 \times 10^{-4}$	$1 - 6,73 \times 10^{-5}$
<i>I-NP_{C01 MAX}</i>	$1 - 3,29 \times 10^{-4}$	$1 - 1,99 \times 10^{-3}$	$1 - 7,93 \times 10^{-4}$

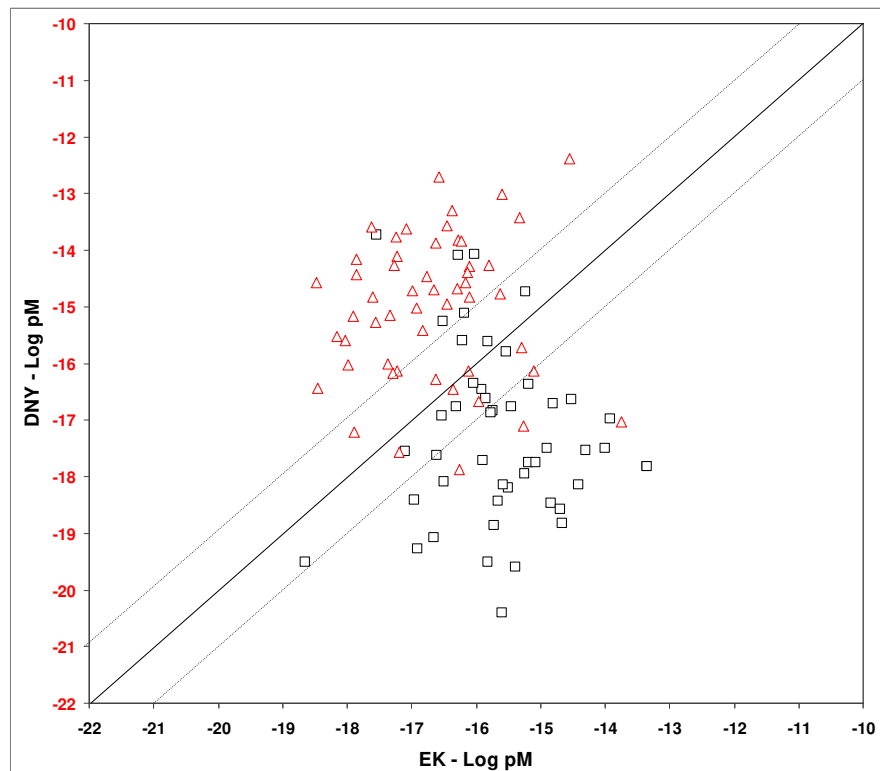
3.14 Az északkeleti (EK) és délnyugati (DNY) allélgyakorisági adatbázisaink genetikai struktúrájának vizsgálata

A korábbi vizsgálatok alapján kijelenthető, hogy a teljes (T) populációs adatbázis inhomogénnek tekinthető, amelynek mértéke befolyásolja a géndiverzitást és az individualizáló erőt. Az EK és a DNY populációk strukturáltságának vizsgálatát a populációt alkotó egyedek vizsgálatával folytattuk. Első lépésben azt a kérdést vizsgáltuk meg, hogy az elemzett EK és DNY populációs adatbázisok hatással vannak-e a genetikai profilok gyakoriságára.

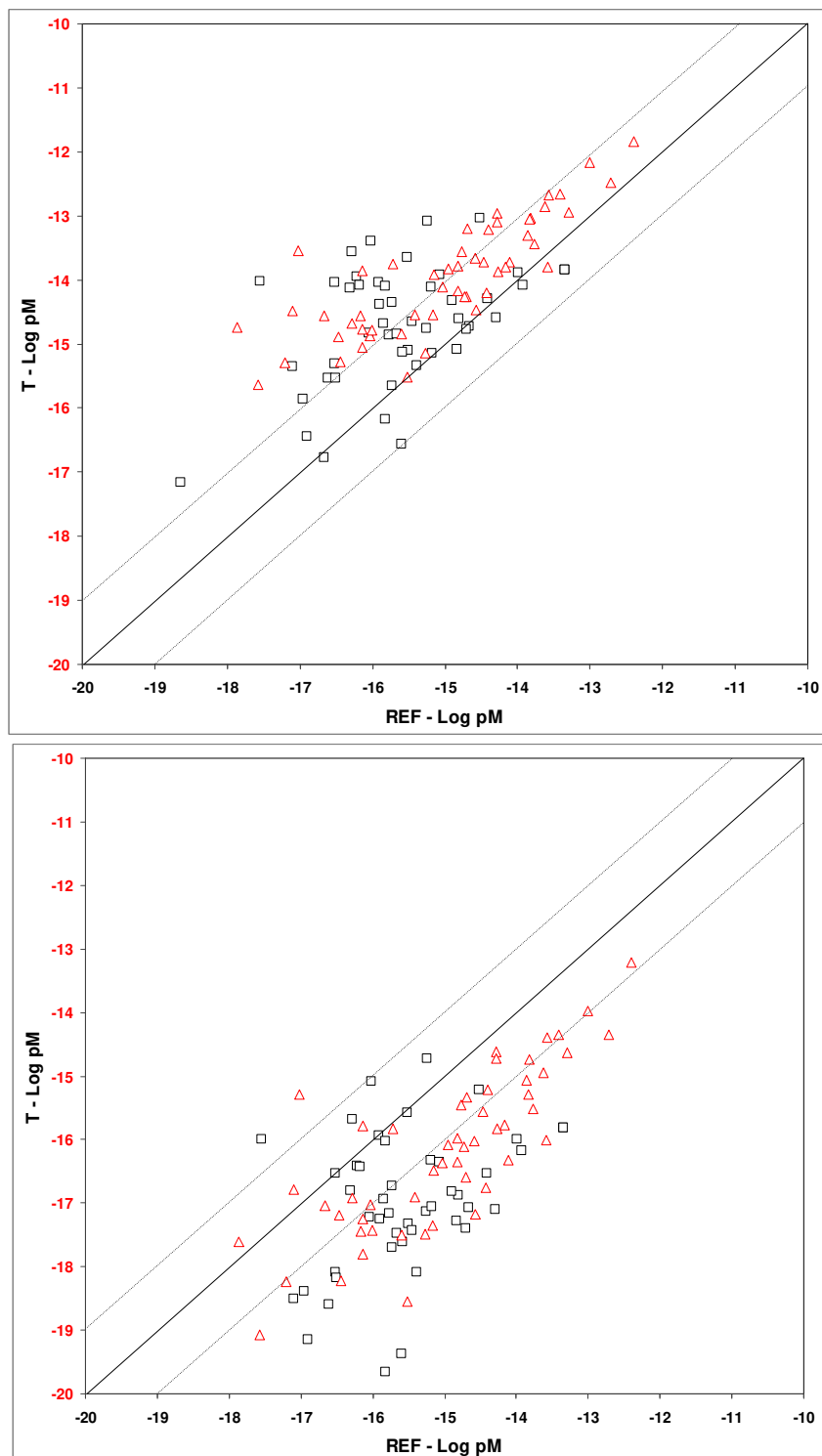
A 6. ábrán lévő „scatter-plot” diagramon jól kivehető, hogy az EK-DNY referencia adatbázis cseréje hatással van a genotípus gyakoriságokra, ugyanis a diagram pontjait az

átló két csoportra osztja fel. Ez alapján véve nem meglepő, hiszen a korábbi eredmények alapján a populációk között szignifikáns alléloszlás különbséget mutattunk ki ($F_{ST} = 0,03398$; $p = 0,000$). Ha az egyedek többségénél a populációkban nincs lényegi – több nagyságrendnyi – különbség a profilgyakorisági ($P \sim pM$) értékek között akkor a diagramon egy csoportba rendeződött pontthalmazt kellene látnunk. Az ábrán jól látható, hogy a két populáció közötti P érték különbség többnyire ugyanabba a populációba sorolja az egyedek jó részét. Az egyedek többségének a referencia populációban nagyobb a $P(\sim pM)$ értékük mintha azt a másik adatbázis allélgyakorisági értékeiből kalkuláltuk volna. Vannak azonban olyan egyedek, amelyek a $P(\sim pM)$ értékpárjuk alapján nem a referencia populációjukba sorolnak. Mindkét populációban találhatóak olyan egyedek is, amelyekre nincs hatással az adatbázis csere, tehát az átló vagy annak $\pm 10 \times P(\sim pM)$ tartományú körzetében csoportosulnak (5. ábra).

5. ábra A *DeerPlex I-II* genotípusokra kalkulált sörét diagram („scatter plot”). Az EK (fekete négyyszög) és a DNY (piros háromszög) populációs adatbázist alkotó egyedek *EK*- észak-keleti és *DNY*- dél-nyugati allélgyakorisági adatbázis alapján kalkulált $\log_{10}(pM)$ értékpárjai.



6. ábra. A *DeerPlex I-II* genotípusokra kalkulált sörét diagramok („scatter plot”). Az északkeleti (EK) és a délnyugati (DNY) populációs adatbázist alkotó egyedek saját (*REF* - referencia) populációs adatbázisa alapján, és a (*T*) teljes allélgyakorisági adatbázis alapján kalkulált $\log_{10}(pM)$ értékpárjai, az A-diagram (fent): $F_{ST} > 0$; B-diagram (lent): $F_{ST} = 0$ esetén.



A 6. ábrán azt a helyzetet szimuláltuk, amikor a célprofillal véletlenszerűen egyező egyed valószínűségét számítjuk ki úgy, hogy adatbázisként nem a referencia-, hanem a teljes populáció allélgyakorisági értékeit használjuk. Az ábra jól szemlélteti, hogy a referencia adatbázisok használata során a random egyezési valószínűségi érték a szubdivíziós paramétertől (F_{ST}) függően milyen mértékben tért el a teljes adatbázis által kalkulált értéktől. Az A diagram pontjainak x és y koordinátáinak kiszámításakor a teljes populációban a lokuszonként becsült F_{ST} értékeket használtuk, a B diagramon ezt a paramétert 0-ra állítottuk. Mind a két populáció egyedeinek profilgyakorisága - többségében - csak akkor nagyobb a saját (referencia) populációjában, ha a teljes populációban nem korrigáljuk a gyakorisági értékeket az F_{ST} paraméterrel. Ennek oka az, hogy az EK és DNY tagoltságát nem ismerjük és ebből következően az EK és DNY populációkon belüli várt értékeket (H_E) nem korrigálhatjuk a szubdivíziós paraméterrel. A teljes populációban az F_{ST} paraméterrel elvégzett korrekció következményeként a várt heterozigóta gyakoriságok alacsonyabbak a homozigóták magasabbak lesznek, mint az $F_{ST}=0$ esetben, tehát a profilgyakorisági értékek növekednek. Ennek eredményeképpen – ami az A diagramon is látható - az EK és DNY populáció egyedei nem a referencia, hanem a teljes adatbázisba sorolnak. Ha meg tudnánk becsülni az alpopulációk struktúráltságának mértékét és azzal kompenzálnánk a genotípusok várt gyakoriságát valószínűleg a teljes adatbázis alapján kalkulált nem korrigált értékektől minden esetben magasabb genotípus (profil) gyakoriságot kapnánk. Ennek érdekében az alpopulációk összetételének további vizsgálatára van szükség.

Az 5. ábrán lévő scatter-plot diagram jól szemlélteti, hogy EK és a DNY populáció bizonyos egyedei P értékpárjuk alapján nem a referencia populációba sorolnak. Felvetődik tehát annak lehetősége, hogy egyik alpopuláció (EK és DNY) sem tekinthető genetikailag homogénnek. Ennek pontosabb vizsgálatához elvégeztük minden adatbázisban szereplő profil konfidencia intervallumának becslését. Az egyedek – mindkét adatbázis szerint kalkulált – profil gyakoriságát és azok 95%-os konfidencia intervallumának logaritmusát a közöttük lévő különbség nagysága szerint csökkenő sorrendben, az F.3-a., F.3-b. ábrákon tüntettük fel. A becsült gyakoriságok között több nagyságrendnyi különbséget mutattunk ki, amelynek mértéke jellemzi ez adott egyed „populációs hűségét”. Ez a gyakoriságbeli különbség adja az ún *LBM* („*likelihood based method*”) módszer alapját (Yang et al., 2005), amelynek értelmében az egyedek

„populációs hűségét” egy valószínűségi hányadossal (LR_{LBM}) jellemezhetjük. Ennek számlálója és nevezője a két, kérdéses forrás populációban becsült genotípus gyakorisága. A legnagyobb LR_{LBM} értékkel rendelkező egyed az EK populációban az EK40, amely szerint ez a gímszarvas kb. 64500-szor valószínűbb, hogy a saját populációjából és nem a DNY populációból származik. A DNY populációban ugyanez az érték a DNY41 egyednél a legmagasabb, a referencia populációjából származás esélye kb. 10000-szer nagyobbra bizonyult. Tehát a forrás populációból származás esélye, mindkét egyednél több mint 99,99%.

A minták egyedszámából becsült genotípus gyakoriságok bizonytalanságát azok konfidencia intervallumával jellemezhetjük. Azoknak az egyedeknek a száma, amelyek mindkét populációs adatbázis (EK-DNY) alapján becsült profilgyakoriságának konfidencia intervalluma nem fedett át, mindkét populációban az egyedek csak kb. 44%-a (EK: 21/48; DNY: 23/52). Ebből következően a gímszarvas egyedek ún. *LBM* módszeren alapuló populációs eredetének valószínűsítése nem adna megbízható eredményt. Ha tehát a populációs adatbázis nagyságával (EK- 48, DNY- 52) is korrigálunk, vagyis a konfidencia intervallumok két szélső értéke közötti legkisebb különbséggel kalkuláltunk akkor az EK40 egyed esetében a profilgyakoriság különbség már kb. 195-szeresére, a DNY41 egyednél pedig hozzávetőlegesen 55-szörösére csökkent. A kis populációból adódó mintavételi hiba következtében a referencia populációból származás esélye az EK40 egyednél még mindig magas kb. 99,5%, a DNY41 egyednél pedig 98,2%.

Mint az a „scatter plot” diagram eredményéből is látható, mindkét populációban előfordulnak olyan egyedek is, amelyek több lokuszos profiljának gyakorisága nem a referencia populációban a legmagasabb. Ezek közül azonban csak egy-egy egyednél (EK42, DNY45) nem fednek át a profilgyakoriság konfidencia intervallumai. Az EK egyedek 16,6%-ában (8/48) volt magasabb a DNY-i adatbázis alapján kalkulált profilgyakorisági érték, mint a referencia populáció szerint számított P érték. Hozzávetőlegesen ugyanezt az arányt kaptuk a DNY egyedek esetében (17,3% - 9/52) is. Az egyedek profilgyakorisági értékeinek vizsgálata alapján mindkét alpopuláció további genetikai tagoltsága valószínűsíthető. Érdekes tény, hogy az északkeleti populációban a populációjukhoz kevésbé „hű” egyedek ($P_{EK} < P_{DNY}$) 75%-a Valkó (Gödöllői-dombság) környékéről származik.

Az ún. *LBM* („*likelihood based method*”) módszert elvégeztük a populáció minden egyedénél, úgy hogy *A* a referencia, *B* a teljes adatbázis volt. Az EK populáció egyedeinek mindössze 8%-nál (4/48), a DNY egyedek kb. 6%-nál (3/52) nem fedett át a 10 lokuszos profilgyakoriság konfidencia intervalluma, tehát a profilgyakoriság és ennek következtében a *pM* érték számításánál megbízható eredményt kaphatunk abban az esetben is, ha azt nem a referencia adatbázisok alapján becsüljük.

3.14 Esettanulmány

A *DeerPlex I-II* rendszerünket egy bűnügyben is alkalmaztuk már, amelyben egy ismert személy ellen nagyobb értékű lopás büntettének elkövetése miatt indult eljárás. Az ügy orrvadászati eset volt, azonban az akkori jogrendszerben még nem létezett ilyen büntető tétel. A feltételezett elkövető megalapozottan volt gyanúsítható azzal, hogy két gímszarvas bikát illegálisan ejtett el. Az ügynek volt egy szemtanúja is akit - a tanúvallomás alapján - a gyanúsított személy arra kényszerített, hogy az elejtett állatok testrészeit (köztük a levágott agancsos fejeket is) a helyszínről a gyanúsított birtokában lévő hűtőházba szállítsa. A hatóság az eljárás során a tett helyszínéről egy levágott csánkot biztosított. A helyszínelők a gyanúsított hűtőházában megtalált levágott szarvas fejeket lefoglalta és bűnjelként a szakértő részére bocsátotta. A tanú munkásruháját is beküldték szakértői vizsgálatra.

Mintát vettünk a két agancsos fej izomszövetéből, a tanú tulajdonában lévő munkásnadrágon kimutatott vérszennyeződésből és a helyszínen lefoglalt szarvas csánk izomszövetéből. A DNS preparálási eljárás, a PCR reakciók, az elektroforézis és a fragmensméret meghatározás körülményei megegyeztek az 2. fejezetben leírt módszerekkel.

A vizsgálatok eredményeképpen a munkásnadrág vérszennyeződéséből, a tett helyszínén rögzített csánk izomszövetéből valamint az egyik szarvasbika fejének izomszövetéből a 12. táblázatban közölt 10 lokuszos genetikai profilokat mutattuk ki.

Az alléltípusok meghatározása manuális módon történt, ismert genotípusú referencia szarvas egyed alkalmazása mellett. Az 1. szarvasbikából kimutatott T172- 6-os és a C01- 28-as allélokat a populációs felmérések során nem detektáltuk, ezek új alléloknak bizonyultak (12. táblázat).

12. táblázat. A bűnjelmintákból kimutatott *DeerPlex I-II* genetikai profilok.

	<i>C229</i>	<i>T26</i>	<i>T193</i>	<i>T108</i>	<i>T123</i>	<i>T501</i>	<i>T172</i>	<i>T156</i>	<i>C01</i>	<i>T507</i>
1. szarvasbika	15	14/ 15.1	10/ 14	14	12/ 16	7/ 11.2	5/ 6	9/ 13	20/ 28	7/ 8
2. szarvasbika	15	8/ 12	16/ 20	7/ 10	14/ 16	7.1/ 10.2	5/ 13	9/ 13.2	25/ 30	7/ 8
csánk	15	8/ 12	16/ 20	7/ 10	14/ 16	7.1/ 10.2	5/ 13	9/ 13.2	25/ 30	7/ 8
nadrág	15	8/ 12	16/ 20	7/ 10	14/ 16	7.1/ 10.2	5/ 13	9/ 13.2	25/ 30	7/ 8

A 2. szarvasbika alléljai minden lokuszon megegyeztek a nadrág vérszenyeyződéséből és a helyszínen rögzített csánkból kimutatott *DeerPlex I-II* genetikai profillal. Feltételezhető volt tehát, hogy a szemtanú nadrágján lévő vérszenyeyződés és a csánk ugyanattól a 2. szarvasbikától származott. Kiszámoltuk annak a feltevésnek a valószínűségét (mint a védelem általánosan a szakértő felé idézett kérdése), hogy a nadrágon lévő vérszenyeyződés és a csánk nem a 2. szarvasbikától, hanem egy azzal nem rokon másik véletlenszerűen ugyanazzal a *DeerPlex I-II* genetikai profillal rendelkező szarvastól származik. A vád és a védelem hipotézisének hányadosa (*LR*) adja bizonyíték statisztikai erejét. Ha a laboratóriumi vagy genotipizálási hiba lehetőségét kizárjuk és feltételezzük a lokuszok függetlenségét (vagyis a C01 marker genotípustól eltekintünk), akkor az *1/LR* érték a profilgyakorisággal megegyezik. A teljes populációs adatbázis alapján, a minimális allélfrekvencia érték fölötti gyakoriságokat alkalmazva a 2. szarvasbika 9 lokuszos *DeerPlex I-II* genetikai profiljának gyakorisága $1,516 \times 10^{-13}$. Ennek alapján annak a várható esélye, hogy a 2. szarvasbikán kívül másik egyed is rendelkezhet a kimutatott 9 lokuszos *DeerPlex I-II* genetikai profillal több, mint 1 a 10 billióhoz. A vád hipotézise tehát több mint 10 billiószor valószínűbb, mint a védelem feltevése. A *DeerPlex I-II* segítségével megoldott orvvadászati eset jól példázza annak rendkívül magas identifikáló erejét és igazságügyi genetikai alkalmazhatóságát.

Az eset az észak-magyarországi régióban történt. Egy olyan eset szimulációját végeztük el, amiben fontos információ lehet a hatóság részére az egyed populációs eredete, vagyis egy gímszarvastól származó genetikai profil (pl. agancsból) alapján beazonosítható lenne-e az egyed populációs származása. Elvégeztük az 1. és a 2. szarvasbika populációs eredetének valószínűsítését 9 lokuszos *DeerPlex I-II* genetikai profiljuk alapján. Az 1. szarvasbika profilgyakoriságainak konfidencia intervallumai

átfedtek ($P_{EK}=6,489 \times 10^{-16}$, *k.i.*: $9,749 \times 10^{-15}/4,319 \times 10^{-17}$; $P_{DNY}=1,623 \times 10^{-17}$, *k.i.*: $3,144 \times 10^{-16}/8,376 \times 10^{-19}$). A 2. szarvasbikánál nem volt átfedés ezen értékek között ($P_{EK}=1,197 \times 10^{-13}$, *k.i.* $1,167 \times 10^{-12}/1,228 \times 10^{-14}$; $P_{DNY}=2,416 \times 10^{-16}$, *k.i.*: $4,077 \times 10^{-15}/1,432 \times 10^{-17}$). A kalkulált LR_{LBM} az 1. szarvasbikánál 39,98 a 2. szarvasbika esetében 495,5. Mindkét egyednél igazolható az észak-magyarországi eredet azonban figyelembe kell vennünk, hogy az EK és DNY adatbázis nem homogén, valamint adatbázisunk sok magyarországi élőhelyet nem reprezentál. Ennek következtében ezek az adatok csak tájékoztató jellegűek, ám jól szemléltetik, hogy további populációs minták analízisével a *DeerPlex I-III* alkalmas lehet a gímszarvas egyedek populációs eredetének meghatározására is.

4 Új megállapítások és következtetések

Magyarországon elsőként végeztük el polimorf gímszarvas tetranukleotid mikroszatellita (STR) lokuszok igazságügyi genetikai vizsgálatát. A *DeerPlex I-II* marker készlettel létrehoztunk egy olyan vizsgálati rendszert, ami hozzájárult a magyarországi vadvilág igazságügyi genetika (az ún. *wildlife forensic genetics*) fejlődéséhez mind a laboratóriumi technológia szintjén, mind elméleti, tudományos értelemben. Továbbá reméljük, hogy a *DeerPlex I-II* gímszarvas mikroszatellita készlet kifejlesztése irányadó lehet más fajokra – pl. őz és dámszarvas – tervezendő, egyedazonosítást célzó vizsgálati rendszer létrehozásához.

4.1 Gímszarvas tetranukleotid mikroszatellita lokuszok

A markerek PCR primereinek tervezéséhez feltételeztük, hogy különböző *Cervidae* fajok ortológ mikroszatellita génjeinek határoló (ún. flanking) régióiban szekvencia eltérések (mutációk) lehetnek. A primer kötőhelyekben bekövetkező pont vagy in/del mutációk genotipizálási hibákat okozhatnak. Ennek következtében a primer oligo és a templát szál nem megfelelő hibridizációja allélkieséshez (un. „drop out”) vagy meglehetősen kiegyensúlyozatlan heterozigóta genotípushoz vezethet (*Butler, 2011; Leibelt and Budowle 2003; Takayama et al., 2007*). A zoo-klónozási eljárással a munkánk kezdetén rendelkezésre álló - őszvérszarvasból (*Odocoileus hemionus*) és amerikai vapitiből (*Cervus canadensis*) izolált - 22 tetranukleotid STR markerből 14 gímszarvas ortológ teljes szekvenciáját sikerült meghatározzuk. A 14 lokusz teljes DNS szekvenciáinak illesztése során a T123 lokusz kivételével minden marker esetében flanking régió mutációkat mutattunk ki. A határoló régiók primer kötőhelyeit vizsgálva négy lokusznál detektáltunk szekvencia eltérést, ami a szekvencia meghatározások nélkül a korábban említett genotipizálási hibákat okozhatta volna. A határoló régió méretéhez viszonyított legtöbb mutációt az őszvérszarvas eredetű C01, C143, C229 valamint a vapitiből izolált T26 STR markerek esetében mutattuk ki. Paleontológiai leletek és filogenetikai számítások alapján a - több recens alfajt is magában foglaló - Nyugati kládba tartozó *Cervus elaphus hippelaphus* és a Keleti kládba sorolt *Cervus canadensis* közös ősének elkülönülése legkorábban 2 millió éve a korai Pleisztocénben

kezdődhetett. A két kládot magában foglaló *Muntiacinae/Cervinae* csoport és az *Odocoileinae/Alcinae/Capreolinae* fejlődési irány elválása jóval több, mint 7 millió éve következhetett be (Pitra et al., 2004, Ludt et al., 2004). A filogenetikai eredmények alapján a C01, C143, C229 markerek nagyszámú szekvencia eltérése a gímszarvas - őszvérszarvas ortológok között nem meglepő, hiszen a két faj evolúciós értelemben távolabb áll egymástól, mint a vapiti a gímszarvastól, így a felhalmozódott mutációk száma is magasabb. A T26 marker flanking régiójának nagyfokú szekvencia-, és a későbbiekben tárgyalt allél polimorfizmusa arra enged következtetni, hogy evolúciós tekintetben a vizsgált mikroszatellita gének között a legősibb lehet. A mikroszatellita evolúciósan korábbi keletkezését a szarvasmarha ortológgal történt szekvencia összehasonlítások eredményei is jól mutatják. Ennél a markernél tapasztaltuk a legalacsonyabb fokú szekvencia hasonlóságot, ami 79,37% volt.

4.2 Polimorf gímszarvas tetranukleotid mikroszatellita markerek

A polimorfizmus tesztek és a populációs felmérések során a 14 gímszarvas ortológból 10 polimorf marker bizonyult alkalmasnak az egyedek elkülönítésére alkalmas marker készlet létrehozásához (F.2. táblázat). A teljes populációs felmérés eredményeképpen minden lolusznak meghatároztuk a populáción belüli alléeloszlását. Ennek során 135 allélt, ezen belül 28 köztes allélt különítettünk el. 50 allél DNS szekvencia meghatározását végeztük el, amelyből 10 intermedier típusú volt. A mikroszatellita allélok szekvencia vizsgálatával megteremtettük a lehetőségét a nemzetközileg elfogadott STR allél nevezéktan alkalmazásának és igazoltuk - a T26, T501, T507 és a T156 lokuszok esetében detektált - egy, két vagy három bázis egységgel eltérő un. intermedier típusú (köztes) allélok jelenlétét és azok korrekt elkülönítését (F.4-a., F.4-b. táblázatok).

A legkevesebb, 5 alléllal a C229 rendelkezett, a legpolimorfabbnak a T26 és a T501 markerek (21 és 23 allél) bizonyultak. A T501, Jones és mtsai által elvégzett, 43 vapiti egyed alapján végzett felmérésben is a legváltozékonyabb volt, összesen 11 db T501 allélt azonosítottak. A T501 nagy allélszáma az ismétlődő régiójában lévő A₁₀ homopolimer ismétlődéseinek köszönhető. Az ismétlődő régió szerkezete alapján a T108, T156, T172 és a T507 egyszerű, a C229, T123, T193 összetett, míg a C01, T26 és

a T501 komplex felépítésű mikroszatellita markerek. A stabil és a polimorf ismétlődő egységek valamint az allélszám alapján a T26 és a T501 hiperpolimorf STR lokuszok.

4.3 Genotipizálás *DeerPlex I-II* markerekkel

Megvalósítottuk a tíz polimorf tetranukeotid STR markert két 5-5 lokuszt tartalmazó multiplex PCR rendszerbe (5-Plex) illesztését. A genotipizáláshoz 5 színes elektroforézis rendszert hoztunk létre. A fluorofór festékeket a rendelkezésünkre álló DS-33 (Applied Biosystems) spektrális kalibráló rendszeréből állítottuk össze. Két olyan 5-Plex rendszert hoztunk létre, amelyben egy-egy pár STR marker rendelkezik ugyanazzal a fluorofórral és a markerek alléltartományai között legalább 20 nukleotid különbség van. A legnagyobb allél mérete nem haladja meg a 250 bázispárt, ami elengedhetetlen a degradált DNS tartalmú minták genotipizálásához. A kisméretű fragmensek preferenciális amplifikációja következtében ugyanis a PCR amplifikáció során a nagyobb (>250 bázispár) amplikont eredményező lokuszok esetében az allél vagy a teljes genotípus kiesésének valószínűsége növekszik (*Butler, 2011*). A két multiplex rendszer alkalmas lehet ásatag csontminták analíziséhez is.

A minták elektroforézise során detektált nem allélikus eredetű termékek (artifaktumok) relatív mennyiségét a multiplex PCR reakciók optimalizált dNTP, Mg^{2+} és primerkoncentrációjának valamint anneálási hőmérsékletének eredményeképpen sikeresen csökkentettük. A T172 és a T501 lokuszok allélcsúcsai mellett detektált artifaktumok („double peak”, „shoulder peak”) csökkentésének érdekében primer módosításra is szükség volt („pig-tailing”). A modifikáció hatására a T172 alléljainál csak egy „peak” volt detektálható, a T501 markerkernél a hármass csúcs jelenségét megszüntettük azonban egy shoulder csúcs továbbra is kimutatható volt minden elektroferogramon. A multiplex PCR reakciók során keletkező nem STR eredetű amplikonok és a T501 markernél detektálható shoulder-peak nem teszi lehetővé a kevert DNS tartalmú minták analízisét, az erre irányuló fejlesztésre továbbra is szükség van.

A fragmensméret meghatározáshoz a 2. fejezetben részletesen ismertetett ABI PRISM 3130xl kapilláris elektroforézis rendszert használtunk. A GeneScanTM 500 LIZ (Applied Biosystems) belső méretstandard alkalmazásával a populációs felmérés során kalkulált allélcsúcs méreteket 1 nukleotid kiterjedésű és maximálisan 0,25 nt szórású

tartományokba csoportosítottuk. A *DeerPlex I-II* lokuszokkal végzett genotipizálás un. manuális módját valósítottuk meg. Első lépésben az egyedhez tartozó két multiplex PCR reakció során felsokszorozott allélikus eredetű DNS-szakaszok méretének meghatározására került sor az ABI PRISM 3130 kapilláris-gélelektroforézis készülék és a GeneMapper® ID ver3.2 szoftveren végzett analízis alkalmazásával (ld 2. fejezet). Az ismeretlen allélok típusát az határozza meg, hogy az allélcsúcs mérete melyik - a populációs vizsgálatok során detektált - allélhoz tartozó, átlaggal és 0,25 nt szórással jellemezhető, maximálisan 1 nt terjedelmű mérettartományba illeszthető be és a kimutatott allél fragmens-mérete hogyan viszonyul a fragmensméret meghatározás során használt standard allélokhöz.

A genotipizálásnak ez a módja igen érzékeny az elektroforézis körülményeinek változására, amelyek az egyes alléltípusok téves meghatározását eredményezhetik. A probléma kiküszöbölésére az igazságügyi vizsgálatok során alléllétrát alkalmaznak (Butler, 2011). Ennek hiányában minden elektroforézishez standardként egy olyan referencia gímszarvas tehén *DeerPlex I-II* profilját használtunk, amelynek minden alléljának mérete és DNS szekvenciája is ismert volt.

4.4 A *DeerPlex I-II* rendszerek fajspecifikusságának vizsgálata

A *DeerPlex I-II* markerek PCR primereinek fajspecifikusságát - vagyis a primerek keresztreakcióját - az emberen, és a leggyakoribb magyarországi párosujjú patás fajokon is elvégeztük. *DeerPlex I-II*-vel végzett PCR reakciók során az ember, a vadsertés és az őz mintákban ampikonokat nem mutattunk ki. A muflon mintából három, a szarvasmarha mintából az elektroferogramon egy - nem allélikus mérettartományba eső - csúcsot detektáltunk. A PCR primerek keresztreakáló képessége tehát a *Cervidae* család tagjain kívül igen alacsony mértékűnek tekinthető, a gímszarvas profilok értékelését még kevert minta esetében sem zavarhatja. Ennek értelmében a vizsgálatot dámszarvas mintán folytattuk, amelynek során 9 lokusz mutatott keresztreakciót. Az allélcsúcsok morfológiája alapján 10 dámszarvas STR allélt detektáltunk, amelyből 5 fragmens mérete volt beilleszthető a gímszarvasnál detektált allélikus pozíciókba. A T26 marker alléltartományában PCR terméket nem mutattunk ki. A gímszarvas populációs minták genetikai profiljaiban a T26 marker kiesését nem

észleltük, ezért joggal feltételezzük, hogy a minták közé dámszarvas egyed nem került. A vizsgálatok eredményeképpen kijelenthető, hogy a gímszarvas genotípusokat – így a populációs mintákat - egyéb patás vadfajok (őz, muflon, vaddisznó, dámszarvas) és a mintavételezésben részt vevő hivatásos vadászok nem kontaminálhatták. A *DeerPlex I-II* a jelen dolgozatban részletezett PCR, elektroforézis és fragmens méret meghatározás körülményeit alkalmazva a *Cervinae* alcsaládon belül a dámszarvasra nem mutatható ki teljes, 10 lokuszos genetikai profilt. Az alcsalád mitokondriális citokróm b génjének filogenetikai vizsgálata alapján az amerikai vapitivel együtt külön filogenetikai csoportba (Keleti kládba) tartozó szikaszarvas szekvenciálisan igen hasonló lehet, ezért ezen a szarvasfajon végzett fajspecifikussági teszt indokolt lehet. Állati eredetű biológiai anyagmaradványok vizsgálatánál problémát okozhat, hogy egy mintán belül nemcsak egyedek-, hanem közelrokon fajok keveredése is előfordulhat, amely gondot okozhat a kevert profilok kiértékelése során. Vadfajok egyedazonosítását megelőzően – faj/alfajspecifikus inszerciós/deléciós mutációkra tervezett real-time PCR eljárást alkalmazva - indokolt lehet az adott minta fajspektrumának kvalitatív és kvantitatív vizsgálata.

4.5 A *DeerPlex I-II* markerek függetlenségének és kromoszómás lokalizációjának vizsgálata

Igazságügyi genetikában alkalmazott mikroszatellita markerek kromoszómás helyének pontos ismerete elengedhetetlen, hiszen a profilgyakoriság számításának alapfeltétele az alkalmazott markerek független öröklődése. A Slate-féle kapcsoltsági térkép alapján marker készletünk 4 tagjának lokalizációja már ismert volt (Slate *et al.*, 2002). Az STR mikroszatellitáink kapcsoltsági viszonyait tisztáznunk kellett. Kétféle megközelítést alkalmaztunk. Populációstatisztikai számításokat végeztünk minden lehetséges lokuszpár függetlenségének vizsgálatára, amelyhez a *LD* tesztet alkalmaztuk. A kapcsoltsági teszt feltétele a Hardy-Weinberg egyensúly, ugyanis a populáción belül a pánmiktikus állapot eltérése két független lokusz kapcsoltsági egyensúlyát szignifikánsan elvethetővé teheti (Hartl and Clark, 2007). Megvizsgáltuk a 10 marker lokuszon belüli allélikus egyensúlyi állapotát minden populációban. A szignifikancia szint Bonferroni korrekcióját alkalmazva a tíz lokusznál csak a T501 markernél a teljes

és az EK populációban találtunk szignifikáns eltérést a HW egyensúlytól. A markereket egyenként vizsgálva, az 5 % szignifikancia szinten a T501 marker alléljainál minden populációs mintában, a C01 markernél az EK és a T108 lokusznál a teljes populációs mintán belül volt elvethető az a hipotézis, hogy az allélok a lokuszon belül random módon párosodnak. A lokuszon belüli egyensúlyi állapot tehát nem minden marker esetében teljesült, ami befolyással lehetett az ezeket tartalmazó lokuszpárok függetlenségi vizsgálatára. A C01-T501 mind a teljes (Bonferroni korrekció; $\alpha = 0,0011$) mind a délnyugati populációban ($\alpha = 0,05$) az *LD* állapotát mutatta, a T108-C01 pár is szignifikáns eltérést mutatott a *LE* állapottól a teljes populációban. Az EK populációs mintában öt, a DNY populációs mintában két lokuszpár, az összesített populációs mintában nyolc lokuszpár esetében volt elvethető a kapcsoltsági egyensúly a 0,05 % szignifikancia szinten. Azonban nem mutattunk ki egyetlen olyan lokuszpárt sem, amelyik mindhárom populációban a kapcsoltsági egyensúlytalanság állapotában lett volna. A függetlenségi vizsgálatok statisztikai oldalról közelítése bizonyos lokuszpárok esetében (pl. T108-C01, C01-T501) nem tette egyértelművé a markerek egymáshoz való viszonyát.

Az általunk felhasznált polimorf STR lokuszok kapcsoltsági viszonyainak vizsgálatát a szarvasmarha – gímszarvas genom szekvenciájának hasonlóságára alapozott térképezéssel folytattuk. Ennek során egyedül a T507 lokusz ortológja nem volt beazonosítható. A T26 és a C01 markerek ortológjai a 16-os szarvasmarha kromoszómán lokalizáltak, amely a Slate-féle összehasonlító térkép szerint a gímszarvas 14-es kapcsoltsági csoportjával homológ, vagyis a nem független öröklődésük feltételezhető. Annak ellenére, hogy a szarvasmarha T26 szekvencia hasonlósága alacsony a gímszarvas ortológjához képest, és szerkezete alapján mint mikroszatellita gén is megkérdőjelezhető, a későbbi vizsgálatok során a lehetségesen kapcsolt T26-C01 párt figyelembe vettük. Ennek bizonyítása csak a teljes genom ismeretében oldható meg, ami a jelenlegi munka írásakor még nem állt rendelkezésünkre.

4.6 A populációk genetikai diverzitása és genetikai összetétele

Magyarországon elsőként kezdtük el a vadon élő gímszarvas állományok genetikai összetételének vizsgálatát mikroszatellita markerekkel. A populációk struktúráltságát és beltenyészetét vizsgáltuk, ugyanis mindkét hatás következménye a várt heterozigóta genotípusok túlbecslése, amennyiben azt a megfigyelt allélgyakoriságok átlagából számítjuk ki. Ebben az esetben a géndiverzitást és a diszkrimináló erőt, ezekkel együtt pedig az identifikáló erőt is túlbecsüljük, hiszen ezek értéke a heterozigóta genotípusok gyakoriságával nő. A két hatás együttes fellépése viszont az identifikáló erő növekedését is okozhatja (Ayres and Overall, 1999). A H , PIC , PD értékeinek maximumát az északkeleti populációban mutattuk ki, vagyis a két populáció közül az utóbb említett tekinthető a genetikailag változékonyságban és egyben az igazságügyi egyedazonosítás hatékonyságának szempontjából a kedvezőbbnek is. Az F -statisztikai eredmények azt mutatták, hogy a két alpopuláció közötti átlagos genetikai különbség (F_{ST}) szignifikáns volt, de más európai gímszarvas állományokhoz viszonyítva nem volt jelentős mértékű (Kuehn et al., 2003, 2004). A HW egyensúlyi állapot eltérése csak a T108 és a C01 esetében volt magyarázható önmagában a populációk közötti viszonylag magas alléloszlás különbséggel, a T501 esetében a DNY populáción belül becsült igen magas beltenyészet volt az oka a pánmiktikus egyensúlytól való eltérésnek. A teljes populáción végzett globális AMOVA eredménye a lokuszokra és a populációkra átlagolt beltenyészet paramétere 0,0373, Wahlund effektusra 0,034 értéket becsültünk. Az emberi populációk belső F_{IS} értékét és a populációk közötti genetikai különbségek mértékét figyelembe véve mindkét paraméter magasnak számít. A humán és a nem-humán igazságügyi genetikában ajánlott a profilgyakoriság (ezzel együtt az LR) közös leszármazási együtthatóval (θ - theta) végzett korrigálása (NRC, 1996; Linacre et al., 2011), amely a gímszarvas teljes adatbázisának használatakor az F_{IT} -nek felel meg. A teljes populációs adatbázisból kalkulált profilgyakoriság, vagyis ennek következményeként az LR számításához az F_{IS} paraméter lokuszokra átlagolt, míg az F_{ST} paraméter lokuszonként becsült értékét használtuk.

Az alpopulációk inhomogenitását az azokat alkotó egyedek „assignment tesztjével” is kimutattuk. Mindkét (EK és DNY) populációs mintában az egyedek 17%-ának profilgyakorisága nem a referencia populációban volt a legmagasabb, vagyis a

referencia adatbázisaink nem tekinthetők homogénnek. A LR_{LBM} alapú populációs eredet becslés a nem-referencia populációba soroló egyedek viszonylag magas aránya (17%), és az EK-DNY profilgyakorisági értékpárok konfidencia intervallumának nagyarányú (44%) átfedése miatt feltételezhetően nem adna megbízható értéket.

Az alpopulációk (démek) kimutatásához és további vizsgálatához a teljes magyarországi állomány szélesebb körű vizsgálatára lenne szükség. A gímszarvas célprofilok LR értékének kalkulációjához a 100 egyedből álló referencia allélgyakorisági adatbázis használatát ajánljuk, az EK és a DNY adatbázisként való alkalmazása, a feltételezhető és nem korrigálható struktúráltság valamint a kis populáció méret miatt torzított becslést adhat.

4.7 A *DeerPlex I-II* egyedek elkülönítésére való alkalmasságának vizsgálata

Célunk a gímszarvas egyedek genetikai úton történő azonosítása volt ezért meg kellett vizsgálnunk a polimorf mikroszatellitákból álló két 5-Plex PCR rendszerünk identifikáló erejét. Ennek kivitelezéséhez két populáció statisztikai számítást alkalmaztunk. A kapcsoltsági vizsgálatok eredményeképpen a *DeerPlex I-II* gímszarvas egyedek elkülönítésére való alkalmasságának kalkulációját 10 gímszarvas STR markerre és a C01 lokuszt kivéve a multiplex rendszerből, 9 markerre is elvégeztük.

A 100 egyed allélgyakorisági értékeiből megbecsültük a *DeerPlex I-II* genotípusok átlagos egyezési/azonosság valószínűségét (pM/PI_{ave}) a teljes populációban. Az API-CALC v1.0 szoftver alkalmazásával a pM/PI_{ave} értéket korrigálni tudtuk az alpopuláción belül mért szignifikáns beltenyészet (F_{IS}) valamint a teljes populációban kimutatott szubdivíziós (F_{ST}) paraméterrel. A két hatás együttese ugyanis megváltoztatja a lokuszonként várt heterozigóta gyakoriság értékeket, ami torzítaná a pM/PI_{ave} valószínűség becslését. Az elvégzett számítások alapján annak valószínűsége, hogy a populációból véletlenszerűen kiválasztott két gímszarvasnak megegyező 10/ 9 lokuszos genetikai profilja legyen $2,6236 \times 10^{-15} / 1,0056 \times 10^{-12}$. Rendkívül alacsony értékek, amelyek – amennyiben a laboratóriumi hiba kizárt - a téves egyed meghatározás valószínűségét is kifejezik.

A *DeerPlex I-II* multiplex rendszerünk identifikáló erejének számszerű kifejezésének másik módszere a teoretikusan „leggyakoribb” genetikai profil *egyediségének* (un.

„*uniqueness*”) becslése. Az egyediség a populáció nagyságának függvényében annak a valószínűségét adja meg, hogy az adott profil egyedi. Ha nem egy adott profil, hanem a teoretikusan „leggyakoribb” profil egyediségével számolunk, akkor a marker készletünk adott populáción belül értelmezett minimális egyediségének valószínűségét kapjuk meg. A „leggyakoribb” profil számításakor figyelembe vettük az adott EK-DNY populációkon belül létező beltenyészet (F_{IS}) és a teljes populáción belül kimutatható inhomogenitás (F_{ST}) mértékét is, ugyanis a két evolúciós hatás együttesen befolyásolja a profilgyakoriságot. A 10 lokuszos egyediség tartalma $1 - 6,73 \times 10^{-5}$, a 9 lokuszt tartalmazó marker készlet egyediségének valószínűségére $1 - 7,93 \times 10^{-4}$ kaptunk. A számításokhoz az európai gím-, és szikaszarvas állomány 2 millióra becsült egyedszámával dolgoztunk.

Mindkét számítás azt igazolja, hogy amennyiben a laboratóriumi hiba kizárt, akár a 10 akár a 9 polimorf mikroszatellitát alkalmazó készlet nagymértékben alkalmas vad populációban élő gímszarvas egyedek genetikai alapú azonosítására és elkülönítésére. Identifikáló ereje az Egyesült Államok kaukázusi népességére kalkulált Identifiler multiplex PCR „*uniquess*” és kombinált *PD* értékéhez hasonló (Butler, 2011).

5 Köszönetnyilvánítás

A dolgozatom elkészítéséhez szükséges laboratóriumi munka döntő hányadát Gödöllőn, a Mezőgazdasági Biotechnológiai Kutatóközpont egykori Fejlődésgenetikai Csoportjában 2006 és 2008 között végeztem. Mindenekelőtt témavezetőmnek, Dr. Orosz László Professzor Úrnak tartozok köszönettel, aki szakmai és anyagi támogatásával biztosította munkám elkészítését. Köszönettel tartozok a Csoport asszisztens tagjainak Péliné Péliné Tóth Magdolnának, Törökné Sánta Csillának, valamint Gálné Szóráth Nellinek kiváló munkájáért és köszönet illeti volt kollégáimat Dr. Borsy Adrienn, Dr. Molnár Andreát, Dr. Gyurján Istvánt és Dr. Stéger Viktort.

Külön megköszönném Dr. Buzás Zsuzsannának a Mezőgazdasági Biotechnológiai Kutatóközpont egykori munkatársának a PhastSystem rendszeren végzett nagyfelbontású poliakrimalid-gélelektroforézisben nyújtott segítségét.

A munka kivitelezéséhez szükséges elméleti és gyakorlati tudás megszerzésében különösen nagy segítséget nyújtottak a Bűnügyi Szakértői és Kutatóintézet egykori munkatársai, Dr. Egyed Balázs és Dr. Pádár Zsolt.

A szakmai együttműködésért köszönetet érdemel a Szent István Egyetem Állatgenetikai és Állattenyésztési tanszékvezetője, Dr. Zöldág László Professzor és a tanszék DNS Laboratóriumának munkatársa, Dr. Zenke Petra.

Hasznos tanácsaival a statisztikai eredmények kiértékelésében nagy segítséget nyújtott az Eötvös Lóránd Tudományegyetem Genetikai tanszékének adjunktusa, Dr. Pásztor Erzsébet.

A munkám kivitelezéséhez elengedhetetlen mintagyűjtés szervezését a Bükk Nemzeti Park egykori igazgatója, Duska József szervezte meg, ezúton szeretném neki megköszönni segítségnyújtását.

A Microsoft Excell Makrók elkészítésében és a többi szoftveres alkalmazásban hálával tartozom Családom két tagjának, nagybátyámnak Keresztes Miklós villamosmérnöknek és testvéremnek Szabolcsi Márton közlekedésmérnöknek. Köszönöm Családom többi tagjának és Barátaimnak is, hogy sok éven keresztül támogattak munkám elvégzésében.

6 Irodalomjegyzék

Az értekezés alapjául szolgáló közlemények

- Szabolcsi Z**, Egyed B, Zenke P, Padar Z, Borsy A, Steger V, Pasztor E, Csanyi S, Buzas Z, Orosz L. Constructing STR multiplexes for individual identification of Hungarian red deer. J For Sci, under submission.
- Szabolcsi Z**, Egyed B, Zenke P, Borsy A, Padar Z, Zoldag L et al. (2008). Genetic identification of red deer using autosomal STR markers. Forensic Sci Int Genet, 1 Suppl: 623-624
- Borsy A**, Podani J, Stéger V, Balla B, Horváth A, Kósa JP, Gyurján IJr, Molnár A, **Szabolcsi Z**, Szabó L, Jakó E, Zomborszky Z, Nagy J, Semsey S, Vellai T, Lakatos P, Orosz L (2009). Identifying novel genes involved in both deer physiological and human pathological osteoporosis. Mol Genet Genomics, 281:301-313

Az értekezéshez tartozó publikációk

- Andreassen R**, Schregel J, Kopatz A, Tobiassen C, Knappskog PM, Hagen SB, Kleven O, Schneider M, Kojola I, Aspi J, Rykov A, Tirronen KF, Danilov PI, Eiken HG (2012) A forensic DNA profiling system for Northern European brown bears (*Ursus arctos*). Forensic Sci Int Genet, 6(6): 798-809
- Applied Biosystems**, AmpFISTR®NGM™ PCR Amplification Kit User Guide (PN 4425511E), (2011)
- Ayres KL**, **Overall ADJ** (1999) Allowing for within-subpopulation inbreeding in forensic match probabilities. Forensic Sci Int, 103: 207-216
- Ayres KL**, **Overall ADJ** (2004) API-CALC 1.0: a computer program for calculating the average probability of identity allowing for substructure, inbreeding and the presence of close relatives. Mol Ecol Notes, 4(2): 315-318
- Balding DJ**, **Nichols RA** (1994) DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands. Forensic Sci Int, 64:125-140

- Balding DJ (1995)** Estimating products in forensic identification using DNA profiles. *J Am Stat Assoc*, 90: 839-844
- Yang BZ, Zhao H, Kranzler HR, Gelernter J (2005)** Characterization of a likelihood based method and effects of markers informativeness in evaluation of admixture and population group assignment. *BMC Genetics*, 6:50
- Barbanera F, Monica Guerrini, Caterina Beccani, Giovanni Forcina, Petros Anayiotos, Panicos Panayides (2012)** Conservation of endemic and threatened wildlife: Molecular forensic DNA against poaching of the Cypriot mouflon (*Ovis orientalis ophion*, Bovidae). *Forensic Sci Int Genet*, 6(5): 671-675
- Bär W, Brinkmann B, Budowle B, Carracedo A, Gill P, Lincoln P, Mayr WR, Olaisen B (1997)** DNA recommendations. Further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. *International Society for Forensic Haemogenetics. Int J Legal Med*, 110(4): 175–176
- Boldogh S (2002)** Jávorszarvas (*Alces alces* (Linnaeus, 1758)) előfordulása a Putnoki-dombságban. *Folia Historico-Naturalia Musei Matraensis*, 26: 351–352
- Bonnet A, Thévenon S, Claro F, Gautier M, Hayes H (2001)** Cytogenetic comparison between Vietnamese sika deer and cattle: R-banded karyotypes and FISH mapping. *Chromosome Res*, 9(8):673-687
- Botstein D, White RL, Skolnick M, Davis RW (1980)** Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet*, 32: 314-331
- Botstein D, White RL, Skolnick M, Dawis RW (1980)** Construction of a Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms. *Am J Hum Genet*, 32: 314-331
- Brenner CH (1998)** Difficulties in the estimation of ethnic affiliation. *Am J Hum Genet*, 62: 1558-1560
- Brinkman TJ, Person DK, Schwartz MK, Pilgrim KL, Colson KE, Hundertmark KJ (2010)** Individual identification of Sitka black-tailed deer (*Odocoileus hemionus sitkensis*) using DNA from fecal pellets. *Conserv Genet Resour*, 2: 115–118
- Brownstein MJ, Carpten JD, Smith JR (1996)** Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. *Biotechniques*, 20(6): 1004-1010

- Budowle B, Chakraborty R, Giusti AM, Eisenberg AJ, Allen RC (1991)** Analysis of the VNTR locus D1S80 by the PCR followed by high-resolution PAGE. *Am J Hum Genet*, 48: 137-144
- Budowle B, Monson KL, Chakraborty R (1996)** Estimating minimum allele frequencies for DNA profile frequency estimates for PCR-based loci. *Int J Legal Med*, 108(4): 173-176
- Buel E, LaFountain M, Schwartz M (2003)** Using resolution calculations to assess changes in capillary electrophoresis run parameters. *J Forensic Sci*, 48(1): 77-79
- Burbaitė L, Csányi S (2010)** Red deer population and harvest changes in Europe. *Acta Zoologica Lituanica*, 20(4): 179-188
- Burke T, Bruford MW (1987)** DNA fingerprinting in birds. *Nature*, 327: 149-152
- Butler JM (2011)** Advanced topics in forensic DNA typing: methodology. San Diego: Academic Press
- Butler JM, Reeder DJ (1997)** Short Tandem Repeat DNA Internet Database, NIST, Standard Reference Database SRD 130: <http://www.cstl.nist.gov/strbase/>
- Buzás Z, Varga L (1995)** Rapid method for separation of microsatellite alleles by the PhastSystem. *PCR Methods Appl*, 4(6): 380-381
- Caniglia R, E. Fabbri, Claudia Greco, Marco Galaverni, Ettore Randi (2010)** Forensic DNA against wildlife poaching: Identification of a serial wolf killing in Italy. *Forensic Sci Int Genet*, 4(5): 334-338
- Caratti S, Rossi L, Sona B, Origlia S, Viara S, Martano G, Torre C, Robino C (2010)** Analysis of 11 tetrameric STRs in wild boars for forensic purposes. *Forensic Sci Int Genet*, 4(5): 339-342
- Chakraborty R. (1992)** Sample size requirements for addressing the population genetic issues of forensic use of DNA typing. *Hum Biology*, 64: 141-159
- Chakraborty R, Srinivasan MR, and Daiger SP (1993)** Evaluation of standard error and confidence interval of estimated multilocus genotype probabilities, and their implications in DNA forensics. *Am J Hum Genet*, 52(1): 60-70
- Clarke LA, Rebelo CS, Gonçalves J, Boavida MG, Jordan P (2001)** PCR amplification introduces errors into mononucleotide and dinucleotide repeat sequences. *Mol Pathol*, 54(5): 351–353

- Cronin** MA, Palmisciano DA, Vyse ER, Cameron DG (1991) Mitochondrial-DNA in wildlife forensic science: species identification of tissues. *Wildl Soc Bull*, 19: 94–105
- Csányi** S, Lehoczki R, Sonkoly K (2010) National Game Management Database of Hungary. *IJSSC*, 1(4): 34-43
- Csányi** S, **Lehoczki** R (2010) Ungulates and their management in Hungary. pp. 291-318 in: Apollonio, M, Andersen R és Putman R. (eds.) *Ungulate management in Europe in the XXI century*. Cambridge University Press, Cambridge
- Dawnay** N, Ogden R, Thorpe RS, Pope LC, Dawson DA, McEwing R (2008) A forensic STR profiling system for the Eurasian badger: A framework for developing profiling systems for wildlife species. *Forensic Sci Int Genet*, 2(1): 47-53
- DeGiorgio** M, Jankovic I, Rosenberg NA (2010) Unbiased estimation of gene diversity in samples containing related individuals: exact variance and arbitrary ploidy. *Genetics*, 186(4): 1367-1387
- Eichmann** C, Berger B, Parson W. (2004) A proposed nomenclature for 15 canine-specific polymorphic STR loci for forensic purposes. *Int J Legal Med*, 118: 249–266
- Ellegren** H (2004) Microsatellites: simple sequences with complex evolution. *Nat Rev Genet*, 5(6):435-445
- Ewett** IW and **Weir** BS (1998) *Interpreting DNA evidence*. Sunderland, Massachusetts: Sinauer Associates
- Excoffier** L, **Slatkin** M (1998) Incorporating genotypes of relatives into a test of linkage disequilibrium. *Am J Hum Genet*, 62: 171–180
- Excoffier** L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, 131: 479-491
- Excoffier** L, Laval G and Schneider S (2006) Arlequin ver 3.1: An Integrated Software Package for Population Genetics Data Analysis. Computational and Molecular Population Genetics Lab (CMPG), Institute of Zoology, University of Berne
- Fazekas** A, Steeves R, Newmaster S (2010) Improving sequencing quality from PCR products containing long mononucleotide repeats. *Biotechniques*, 48(4): 277-285

- Füredi S, Woller J, Pádár Zs, Angyal M. (1999)** Y-STR haplotyping in two Hungarian populations. *Int J Legal Med*, (113): 38-42
- GenBank Home** - National Center for Biotechnology Information
<http://www.ncbi.nlm.nih.gov/genbank/>
- Gill P, Kimpton C, D'Aloja E, Andersen JF, Bar W, Brinkmann B, Holgersson S, Johnsson V, Kloosterman AD, Lareu MV, et al. (1994)** Report of the European DNA profiling group (EDNAP) - towards standardisation of short tandem repeat (STR) loci. *Forensic Sci Int*, 65(1): 51–59
- Gill P, Brinkmann B, d'Aloja E, Andersen J, Bar W, Carracedo A, Dupuy B, Eriksen B, Jangblad M, Johnsson V, Kloosterman AD, Lincoln P, Morling N, Rand S, Sabatier M, Scheithauer R, Schneider P, Vide MC. (1997a)** Considerations from the European DNA profiling group (EDNAP) concerning STR nomenclature. *Forensic Sci Int*, 87(3): 185–192
- Gill P, Sparkes R, Kimpton C (1997b)** Development of guidelines to designate alleles using an STR multiplex system. *Forensic Sci Int*, 89(3): 185-197
- Gill P, Brenner C, Brinkmann B, Budowle B, Carracedo A, Jobling MA, de Knijff P, Kayser M., Krawczak M., Mayr WR, Morling N, Olaisen B, Pascali V, Prinz M, Roewer L, Schneider PM, Sajantila A, Tyler-Smith C (2001)** DNA Commission of the International Society of Forensic Genetics: recommendations on forensic analysis using Y-chromosome STRs. *Forensic Sci Int*, 124(1): 5–10
- Groves C, Grubb P (2011)** Ungulate Taxonomy. The John Hopkins University Press, Baltimore
- Guo SW, Thompson EA (1992)** Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, 48: 361-372
- Guo X, Elston RC (1999)** Linkage information content of polymorphic genetic markers. *Hum Hered*, 49(2): 112-118.
- Gustavsson II, Sundt CO (1968)** Karyotypes in five species of deer (*Alces alces* L., *Capreolus capreolus* L., *Cervus elaphus* L., *Cervus nippon nippon* Temm. and *Dama dama* L.). *Hereditas*, 60(1-2): 233-248
- Gyurján IJr, Molnár A, Borsy A, Stéger V, Hackler LJr, Zomborszky Z, Papp P, Duda E, Deák F, Lakatos P, Puskás LG, Orosz L (2007)** Gene expression dynamics in

- deer antler: mesenchymal differentiation toward chondrogenesis. *Mol Genet Genomics*, 277(3): 221-235
- Hartl DL, Clark AG (2007)** Principles of Population Genetics, Fourth Edition, Sinauer Associates, Sunderland, Massachusetts
- Hartl GB, Willing R, Lang G, Klein F, Köller J (1990)** Genetic variability and differentiation in red deer (*Cervus elaphus* L) of Central Europe. *Genetics Selection Evolution*, (22): 289-306
- Hatályos Jogszabályok Gyűjteménye:** <http://net.jogtar.hu/>
- Hedrick PW (2011)** Genetics of Population, Fourth Edition, Arizona State University, Tempe, Arizona
- Homonnay Z (2004)** Olcsó húsnak drága a leve. *Magyar Vadászlap*. 13.1:14-17
- Jeffreys AJ, Wilson V, Thein SL (1985a)** Hypervariable minisatellite regions in human DNA. *Nature*, 314: 67-72
- Jeffreys AJ, Wilson V, Thein SL (1985b)** Individual-specific "fingerprints" of human DNA. *Nature*, 316: 76-79
- Jobin RM, Patterson D, Zhang Y (2008)** DNA Typing in populations of mule deer for forensic use in the Province of Alberta. *Forensic Sci Int Genet*, 2(3): 190-197
- Jones DA (1972)** Blood samples: probability of discrimination. *J Forensic Sci Soc*, 12: 355-359
- Jones KC, Levine KF, Banks JD (2000)** DNA-based genetic markers in black-tailed and mule deer for forensic applications. *Calif Fish Game*, 86(2): 115-126
- Jones KC, Levine KF and JD Banks (2002)** Characterization of 11 polymorphic tetranucleotide microsatellites for forensic applications in California elk (*Cervus elaphus canadensis*). *Mol Ecol Notes*, 2(4): 425-427
- Kalmár T, Bachrati CZ, Marcsik A, Raskó I (2000)** A simple and efficient method for PCR amplifiable DNA extraction from ancient bones. *Nucleic Acids Res*, 28(12): e67
- Kanthaswamy S, Premasathan A, Ng J, Satkoski J, Goyal V (2012)** Quantitative real-time PCR (qPCR) assay for human-dog-cat species identification and nuclear DNA quantification. *Forensic Sci Int Genet*, 6(2): 290-295

- Leibelt C, Budowle B (2003)** Identification of a D8S1179 primer binding site mutation and the validation of a primer designed to recover null alleles. *Forensic Sci Int*, 133(3): 220-227
- Kászoni Z (1999.november)** A szikaszarvas. *Magyar Vadászlap*
- Kimpton CP, Fisher D, Watson S, Adams M, Urquhart A, Lygo J, Gill P (1994)** Evaluation of an automated DNA profiling system employing multiplex amplification of four tetrameric STR loci. *Int J Legal Med*, 106:302-311
- Kuehn R, Schroeder W, Pirchner F, Rottman O (2003)** Genetic diversity, gene flow and drift in Bavarian red deer populations (*Cervus elaphus*). *Conserv Genet*, 4(2): 157-166
- Kuehn R, Haller H, Schroeder W, Rottmann O (2004)** Genetic roots of the red deer (*Cervus elaphus*) population in Eastern Switzerland. *J Hered*, 95(2):136-43
- Linacre A, Gusmao L, Hecht W, Hellmann AP, Mayr WR, Parson W, Prinz M, Schneider PM, Morling N (2011)** ISFG: Recommendations regarding the use of non-human (animal) DNA in forensic genetic investigations. *Forensic Sci Int Genet*, 5(5): 501-505
- Lincoln PJ, Thomson J (1998)** *Methods in Molecular Biology, Forensic DNA Profiling Protocols*. Humana Press; 1st edition, Vol. 98
- Liu WS, Mariani P, Beattie CW, Alexander LJ, Ponce De León FA (2002)** A radiation hybrid map for the bovine Y Chromosome. *Mamm Genome*, 13(6): 320-326
- Lorenzini R, Cabras P, Fanelli R, Carboni GL (2011)** Wildlife molecular forensics: identification of the Sardinian mouflon using STR profiling and the Bayesian assignment test. *Forensic Sci Int Genet*, 5(4): 345-349
- Ludt CJ, Schroeder W, Rottmann O, Kuehn R (2004)** Mitochondrial DNA phylogeography of red deer (*Cervus elaphus*). *Mol Phylogenet Evol*, 31(3): 1064-1083
- Mahmut H, Masuda R, Onuma M, Takahashi M, Nagata J, Suzuki M, Ohtaishi N. (2002).** Molecular phylogeography of the red deer (*Cervus elaphus*) populations in Xinjiang of China: comparison with other Asian, European, and North American populations. *Zoolog Sci*, 19(4): 485-495

- Meredith EP, Rodzen JA, Levine KF, Banks JD (2004).** Characterization of an additional 14 microsatellite loci in California Elk (*Cervus elaphus*) for use in forensic and population applications. *Conserv Genet*, 6(1): 151-153
- Molnár A, Gyurján I, Korpos E, Borsy A, Stéger V, Buzás Z, Kiss I, Zomborszky Z, Papp P, Deák F, Orosz L (2007)** Identification of differentially expressed genes in the developing antler of red deer *Cervus elaphus*. *Mol Genet Genomics*, 277: 237-248
- Myers P, Espinosa R, Parr CS, Jones T, Hammond GS and Dewey TA (2013)** The Animal Diversity Web (online). Accessed at <http://animaldiversity.org>
- Nagy E (2002)** A hazai gímszarvas minőségének helyzete, nemzetközi és hazai megítélése. A vadgazdálkodás időszerű kérdései 1., Gímszarvas. Budapest : Dénes Natúr Kiadó: 6-9
- Nei M (1973)** Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci USA*, 70: 3321-3323
- Nelson S (1978)** Nomograph for samples having zero defectives. *Jour Qual Technol*, 10: 42–43
- NRC - National Research Council (1996)** The evaluation of forensic DNA evidence. National Academy Press Washington, D.C.
- Ogden R, Dawnay N, McEwing R (2009)** Wildlife DNA forensics - bridging the gap between conservation genetics and law enforcement. *Endangered Species Research*, 9: 179–195
- Olaisen B, Bar W, Brinkmann B, Budowle B, Carracedo A, Gill P, Lincoln P, Mayr WR, Rand S (1998)** DNA recommendations 1997 of the International Society for Forensic Genetics. *Vox Sang*, 74: 61-63
- Ott J (1999)** Analysis of human genetic linkage. 3rd Edition, John Hopkins University Press
- Paetkau D, Strobeck C (1994)** Microsatellite analysis of genetic variation in black bear populations. *Mol Ecol*, 3(5): 489-495
- Paetkau D, Calvert W, Stirling I, Strobeck C (1995)** Microsatellite analysis of population structure in Canadian polar bears. *Mol Ecol*, 4(3):347-354
- Pamjav H, Triga D, Buzás Z, Vellai T, Lucskai A, Adams B, Reid AP, Burnell A, Griffin C, Glazer I, Klein MG, Fodor A (1999)** Novel application of PhastSystem

- polyacrylamide gel electrophoresis using restriction fragment length polymorphism--internal transcribed spacer patterns of individuals for molecular identification of entomopathogenic nematodes. *Electrophoresis*, 20(6): 1266-1273
- Parson W**, Pegoraro K, Niederstätter H, Föger M, Steinlechner M (2000) Species identification by means of the cytochrome b gene. *Int J Legal Med*, 114(1-2): 23–28
- Pádár Z**, Angyal M, Egyed B, Füredi S, Woller J, Zöldág L, Fekete S (2001) Canine microsatellite polymorphisms as the resolution of an illegal animal death case in a Hungarian zoological gardens. *Int J Legal Med*, 115: 79–81
- Pádár Z**, Egyed B, Kontadakis K, Füredi S, Woller J, Zöldág L, Fekete S (2002) Canine STR analyses in forensic practice: observation of a possible mutation in a dog hair. *Int J Legal Med*, 116: 286–288
- Pechtol J** (2001) Az orvvadászat. Milleniumi vadászati almanach. Budapest: Dénes Natúr Műhely Kiadó. 73-75
- Pérez-Espona S**, Pérez-Barberia FJ, McLeod JE, Jiggins CD, Gordon IJ and Pemberton JM (2008) Landscape features affect gene flow of Scottish Highland red deer (*Cervus elaphus*). *Mol Ecol*, 17: 981–996
- Pitra C**, Fickel J, Meijaard E, Groves PC (2004) Evolution and phylogeny of old world deer. *Mol Phylogenetic Evol*, 33(3): 880-895
- Poetsch M**, Seefeldt S, Maschke M, Lignitz E (2001) Analysis of microsatellite polymorphism in red deer, roe deer, and fallow deer -- possible employment in forensic applications. *Forensic Sci Int*, 116(1): 1-8
- Polziehn RO**, **Strobeck C** (2002) A phylogenetic comparison of red deer and wapiti using mitochondrial DNA. *Mol Phylogenetic Evol*, 22(3): 342-356
- Reiczigel J**, Harnos A, Solymosi N (2010) Biostatisztika nem statisztikusoknak. Javított utánnymás, Pars Kft., Budapest
- Rousset F** (2008) Genepop'007: a complete reimplement of the Genepop software for Windows and Linux. *Mol Ecol Resour*, 8(1): 103-106
- Rychlik W** (2007) OLIGO 7.0 Primer Analysis Software. *Methods Mol Biol*, 402: 35-59
- Slate J**, Van Stijn TC, Anderson RM, McEwan KM, Maqbool NJ, Mathias HC, Bixley MJ, Stevens DR, Molenaar AJ, Beever JE, Galloway SM and Tate ML (2002) A

- Deer (Subfamily Cervinae) Genetic Linkage Map and the Evolution of Ruminant Genomes. *Genetics*, 160: 1587-1597
- Sugár L (2002)** A gímszarvas állomány csökkentés mértéke, üteme és módja. A vadgazdálkodás időszerű kérdései 1. Gímszarvas. Budapest: Dénes Natúr Kiadó. 42-47
- Stéger V, Molnár A, Borsy A, Gyurján I, Szabolcsi Z, Dancs G, Molnár J, Papp P, Nagy J, Puskás L, Barta E, Zomborszky Z, Horn P, Podani J, Semsey S, Lakatos P, Orosz L (2010)** Antler development and coupled osteoporosis in the skeleton of red deer *Cervus elaphus*: expression dynamics for regulatory and effector genes. *Mol Genet Genomics*, 284: 273-287
- Szedzerjei A (1965)** Szarvas. Mezőgazdasági Kiadó, 2nd ed, Budapest
- Takayama T, Takada N, Suzuki R, Nagaoka S, Watanabe Y (2007)** Identification of a rare mutation in a TH01 primer binding site. *Leg Med*, 9(6): 289-292
- Thermo Scientific.** NanoDrop 2000/2000c Spectrophotometer V1.0. User Manual. US patents 6,628,382 and 6,809,826
- Thommasen HV, Thomson MJ, Shutler GG, Kirby LT (1989)** Development of DNA fingerprints for use in wildlife forensic science. *Wildl Soc Bull*, 17: 321–326
- Triga D, Pamjav H, Vellai T, Fodor A, Buzás Z (1999)** Gel electrophoretic restriction fragment length polymorphism analysis of DNA derived from individual nematodes, using the PhastSystem. *Electrophoresis*, 20(6): 1274-1279
- Van Asch B, Alves C, Santos L, Pinheiro R, Pereira F, Gusmão L, Amorim A (2010)** Genetic profiles and sex identification of found-dead wolves determined by the use of an 11-loci PCR multiplex. *Forensic Sci Int Genet*, 4(2):68-72
- Von Köller J, Kabai P, Demeter A (1988)** Untersuchungen zur regionalen Differenzierung ungarischer Rotwildpopulationen anhand morphologischer Geweihmerkmale. *Zeitschrift Für Jagdwissenschaft*, (34): 86-97
- Wada K, Okumura K, Nishibori M, Kikkawa Y, Yokohama M (2010)** The complete mitochondrial genome of the domestic red deer (*Cervus elaphus*) of New Zealand and its phylogenic position within the family Cervidae. *Anim Sci J*, 81(5): 551-557
- Walsh PS, Fildes NJ, and Reynolds R (1996)** Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. *Nucleic Acids Research*, 24: 2807–2812

- Wang DY**, Green RL, Lagacé RE, Oldroyd NJ, Hennessy LK, Mulero JJ (2012) Identification and secondary structure analysis of a region affecting electrophoretic mobility of the STR locus SE33. *Forensic Sci Int Genet*, 6(3): 310-316.
- White NE**, Dawson R, Coghlan ML, Tridico SR, Mawson PR, Haile J, Bunce M. (2012) Application of STR markers in wildlife forensic casework involving Australian black-cockatoos (*Calyptorhynchus* spp.). *Forensic Sci Int Genet*, 6(5): 664-670
- Wictim E**, Kun T, Lindquist C, Malvick J, Vankan D, Sacks B (2013) Developmental validation of DogFiler, a novel multiplex for canine DNA profiling in forensic casework. *Forensic Sci Int Genet*, 7(1): 82-91
- Wright S (1951)** The genetical structure of populations. *Ann Eugen*, 15: 323-354
- Zsolnai A**, Lehoczky I, Gyurmán A, Nagy J, Sugár L, Anton I, Horn P, Magyary I (2009) Development of eight-plex microsatellite PCR for parentage control in deer. *Arch Tierz*, 52(2): 143-149

7 Függelék

Lokusz	Ortológ	GenBank	Forward PCR Primerek (5'- 3')	T_m	Reverz PCR Primerek (5'- 3')	T_m
C229	O. hemionus	AF102257	CCCTGTGGTCTAGCAAATTCATT	61	GGTCTTATCTTAGCTTATACATTGC	61
T108	C. canadensis	AF191798	AATCCATGTTTATCTGCCTTCAGAG	63	GACACAAAGAGTGAGTGGTAAATGAT	63
T26		AF442814	ACTGTCTCCCTCACCAGCCTGGTC	70	CACCATCAGGCAACCGCATCTAGGC	71
T123		AF192395	TTGAGAAGCCCCTGCTGTGTCTT	65	ACAACAGGTAGATAGACTGTCGTTG	64
T193		AF192398	TTCAGCTTCAGCATCAGTCCTTCC	65	CACGTAGAGTAGTAAGTATTCATGAACG	58
T172		AF192397	CTGGGAGATGGAGGTGCTGAAGC	68	GCTAAGTCTCTTCAATTGTGTCCCAAC	67
C01	O. hemionus	AF102248	GAGGTATTTAGGGAAGGGTTGTCTGAT	67	CTGTTTGCATACCAAGGAGAGAGACC	68
T501	C. canadensis	AF442815	TTTACCTGGTAACTCATTTACTCCTC	63	TATAATGTGAGTCAGGAGACTTCTC	63
T156		AF192396	CTAGGCAAGAATACTGGAGTGGGT	65	GGAAAGTTAGACAGAAGCACCAATAC	65
T507		AF442816	AGGCAGATTCTTTACCGCTGAGTC	65	CTACGAGGCAGATACAGTGATG	62
C02		AF442813	ACCTCTCCTCTGGGTTCTCACA	66	CAGAGCTTGAGTGGTGGACTTAATG	66
C143	O. hemionus	AF102252	CAGGCGATCAAGAAATGTGGAATCAATG	66	AGGTGTTGGAGGCTGTGGAGCTGAT	67
C180		AF102255	GGTGGGCATTTCAGTAGATCTTGG	65	GGTCAGAATGCATGGAGGTAATAAGG	66
T107	C. canadensis	AF193019	ATTGACTGGTTTGATCTCCTTGCTAG	61	GAATGAATGAGATACTAATATGGAGAC	61
T27	O. hemionus	AF102244	GTTGGGCTTATGAATCAGGTAGG	63	CTATGGATGCCTTGGTATGTGGTTAA	65
T530	C. canadensis	AF442817	TGTAATGCTTCATTGTGTCTCACAG	65	GGAGTAGGTTGCCATTCCCTTCTC	67
C217	O. hemionus	AF102242	CCTGGCAGGAAGGAGGAGACAG	68	ATCTCCTATCATTATCTGGCCCTTCC	66
C276		AY514441	CTCAGATTTAGAAACAGAACATTCACC	64	TCGTATGTGGTTCAGGGATGTGATC	66
C273		AF102246	TGATGGCATTAAACACGACATTAGAA	61	TTCTTGAAGAATAGAAAGAGCTTG	58
C32	C. canadensis	AY514436	ACTTGCTGCCATCTTGAAGGGACAGG	69	AGGGATTGCGCAGCAACACCTAGATC	69
C105	O. hemionus	AY514440	ACCATATTCAAAGACAGCTCAGAGA	63	TCCTCGTGGAGTTGTAGTTACAG	63
T115	C. canadensis	AF193021	TGCAGCACTTTCACAGCATCATGTT	64	GAGTGGGTTATCATTTCTTCTCC	64

F.1.táblázat 1-3. oszlop: a lokuszok elnevezése, faji eredete, GenBank azonosító száma, 5. és 7. oszlopok: a gímszarvas templátra nem specifikus PCR primerek szekvenciái és olvadási hőmérsékletei (4. és 6. oszlopok).

Lokusz	Genbank	Lokusz (bp)	Flanking régió (bp)		Flanking régió mutációk			Repeat régió	DeerPlex	Dye	Forward/Reverz PCR Primerek	Conc. (nM)	Allélok (bp)
			5'	3'	In/Del	TV	TI						
C229	JN998698	341	190	84	1	4	4	összetett	I	6-FAM	TTATTCATCCACCCATCCATCACCA GGCACATGCTCATAAGTGAAGGGA	6	107-123
T108	JN998701	244	52	128	0	0	I*	egyszerű	I	NED	GTGGAGATAGGTAGACAGATGGAG GCTGATTCAGAGGTTGCCAAAGATG	8	119-159
T26	JN998699	461	121	255	2	6	9	komplex és hipervariábilis	I	6-FAM	CATGCAGATGCTACTCCCTTATGG CCAGAAGAGATACTGCACCCTGAA	11,2	147-213
T123	JN998702	317	30	211	0	0	0	összetett	I	PET	TTGTTCTGTGTCTTAGTGTGTCTGAC CTGAGCTGAGCCACATGATTAT	24	79-135
T193	JN998705	357	154	147	0	0	1	összetett	I	VIC	AGATGATAGACAGATGGGTAGCGG GCTCTACTGCTGTTGTCATCATTACC	16	129-193
T172	JN998704	340	106	190	0	0	0	egyszerű	II	VIC	CCCTTCAACATGTCATTAGAGCC GTTTCTGTCTTTACCACTAGCGCTACCTGG	5,6	105-157
C01	JN998694	479	251	57	3	2	9	komplex	II	NED	GTTCATCCTTCTACCTTACATCCACC CCTCCACACCTGGATCTCAAATAA	5,6	206-258
T501	JN998706	379	168	131	0	I***	5***	komplex és hipervariábilis	II	6-FAM	CTCAGCCTACATGCTTTGACCAAGA GTTTCTGTGACAAGCATATTGTGCAAATCTGACTC	28	142-183
T156	JN998703	258	52	129	0	0	2	egyszerű	II	NED	CCTGGCCTGTGTCTTGAATTGAAC GGCGATGAATACCCAGTCTTGCTCT	7,6	129-185
T507	JN998707	414	111	272	1	0	2	egyszerű	II	PET	GTTATCAGACGACGCTGTGGAGCA TTGAACCCGGGTCTCTTGCAATTGC	24	181-221
C02	JN998695	443	32	368	1	3	4						
C143	JN998696	338	87	219	3	1	6						
C180	JN998697	420	0	386	3	1	2						
T107	JN998700	404	222	133	0	0	2						

F.2.táblázat 1-3. oszlop: a lokuszok elnevezése, GenBank azonosító száma, a lokuszok mérete; 4-9. oszlopok: az 5' és 3' flanking régiók mérete, a gímszarvas és az ortológ szekvenciák eltéréseinek (mutációk) száma, ismétlődő régió struktúrája; 10-14. oszlopok: *DeerPlex* 5-Plex, az alkalmazott fluoreszcens jelölés, primer szekvenciák és azok koncentrációja, *DeerPlex I-II* allélméret tartományok. (In/Del, inszerció/delécio; TV, transzverzió; TI, tranzíció) * Kizárólag a 16-os allél mutatott szekvencia különbséget, egy A/G tranzíciós mutációt, ** A mutációk közé beszámítottuk a 11.2 allélnál detektált A/G tranzíciót és T/G transzverziót.

Lokusz	Flanking régió szekvencia illesztései
C01	GGTAGAGGTTAGAATGGCTCACACAGAAAGAGAATGAAGTGGGAGGTTTCATTGTCCTTCTGTTTCCTTTCAGACCATTGAGATAAGAATTTGGCTGAGACTTCTCTGTATCTGAGGAGTTCAC TAGGGA.....A.....T.....C..... AGGCAC TGTATCAAGTGAGGGTGTGAGATGGGATGTCATCATCTCTTTTGTGGCACATGCTGCAGATTCATTCAGTCTCACCACCTACAATTGTTTCATCCTTCTACCTTACGTCACCT---- RCA.....G.....A.....A.....A.....GTTC CTAACTATTCATCCATTTATTTGAGATCCAA-TGTGGGAGGG-CTGGTTTCTCCAAGG.....G.....G.....
C229	GCTAGAAAAAA-TGTTTTCTCTAGAAAAAATGTCTATATATGTGTTTGAGGTGACATTT CATAAACTGATCATTGTAGCACTATTTATAGTGGTTGAA GTTGCATATAAAAACTGGAACAAACCT T.....A.....T.....T..... ATGACCATTCACTTTTTATTATCCATTTTAAATTATTCATCCACCCATCCATCACCATA R ATCCCTTCACTTATGAGCATGTGCCTATATGTGTGTTTATATGCATACTATGTGCCAGGCATT.....C.....C.....G.....G..... GGTTCTAGGTGCTGGACATGTA
T26	CTTATACTGAAGTTCCAATAGACACGCTCATGATTTAAAGTGTGCAATGCCTATTCAAACACGTATTGTTGTTTTTTTTTCCTCCACCCATGCAGATGCTACTCCCTTATGGGGAAAA R ATAA.....T.....A.....A.....C..... AAGATTCAGGGTGCAGTGTCTCTTCTGGAGACACACCTATAATTGCATCAGAATTACACTATGAAATAATTGTAAGACACAAAGTTAGAGAAGGTAGGAAAACTATGGCAGTAAAAATGCAAAAGAA.....G.....TG.....G.A.....G..... CATGCG-----TAAAAGACAAGTTGAGAATGCTGCCTTGACAGCTTTGACTCTGACCAAGCTAGTCAGGGTAGCTTAGACTATATACAGGCTCAACCAAGGCTAAACTAAGTTCAAT.CAAAAGACATGTA.....T.....C.....-...
T108	TATATATACATGTGGAGATAGGTAGACAGATGGAGAGAGGATAGAGATATAT R GAGATAAATCAGGCAGAGAGAGCAAATACCATCTTTGGCAACCTCTGAATCAGCTACTCAGAATGGAAAAACG..... AGAAAATGTTTGGTATGTAGCATACTAAACCTAAATAAATAAACTTTTAGAGAA
T123	TTGTTCTGTTGCTTAGTTGTGTCTGACTCT R CATAAATCATGTGGGCTCAGCTCAGAGCAAGCCCAATAAACTGATCATAGAGATGTCCAAGGAAACTGAGATATGTCCAGGAATGGGAACACG GGTGGGTGCATCTGACCACTTGATGGTGTGCATCACCTGTGGTTATCACTCTCTGACAACTGCAGCTTCAGACAGACAGAATTATGTCCAAGGGCGGGAAGGGAGGTGGCTAGGAT

Lokusz	Flanking régió szekvencia illesztései
T156	AGCCATTCCCTTCTCCAGGGGATCTTCCTGACCTGTGTCTTGAATTGAACCT <u>R</u> TAGATCAGACAAGACTGGGTATTTCATCGCCCTGCATGTAATGGAAGAAGGATTCTGCTGGTGAATAAAGA <i>G</i> GAGGTTGTATTTTGA AACAGAAAGTCATTTTCAGTTTCTGACTGAGGCTCATGATAAAA <i>G</i>
T172	CCCCAGTTTATACAAGCTAACCTCCATGGAGCTGTGTGGGACAAAATGCGAGATCTTTGATCTAGCATCTCCCTTTCAACATCTCATTTCAGAGCCTAGAACAAATT <u>R</u> TATATGTATGTGGGCTT <i>G</i> CCCAGGTAGCGCTAGTGGTAAAGAACCCACCTGTCAATGCAGGAGATACGGGTTTCGATACTTGGGTTGGGAAGATCCCCTGGAGAAGGGCATGCAACCCACTCCAGTATTCTTCTCTGGAGAATCC CAAGGACAGAGGAGCCTGGCAGGCTATAGTCCATGGGGTTGCAAAGA
T193	TTCAGCTTCAGCATCAGTCCTTCCAGTGAACACCCAGGACTGATCTCCTTTAGGATGGACTGGTTGGATCTCCTTGCAGTCCAAGCCTGCATAAATAATTGCTATATGTTGAGTAGATAGATAGGTG <i>A</i> ATAGACAGATGGGTAGCGGGTAGACAG <u>R</u> CTGAGATTCCTCAGCTCTAAATGAATGGTAATGATGACAACAGCAGTAGAGCTGTGTGATGTTAACAGGTGATATAAACAAGCTCAGGAGCTCAA TGTATAGCCTGATGCACAGTAATCGTTCATGAATACTTACTACTCTACGTG
T501	TGACGGTCCGATGGCTAAGAATCTGCCTTCCAATGCAGCGGGCAAGGGTTCAATGCCTGATCAGGGAAGCTAAGATCCGCAAGCGCTGGGGCAGCTAAGCCTGCTCACTGCTACAACCTACTCAGCCTA <i>T</i> <i>A</i> <i>A</i> CATGCTTTGACCAAGACCCGATGCAACCAATTAAATATATC <u>R</u> TATGAGTCAGATTTGCACAATATGCTTGTACATATTGGGTTTAAATTCTGCCTAGATACCTACAGTTTATTCTCTGAGCCT <i>GG</i> TAGTATCCTCATCCATATAATAACTGTGTCAACTTCACAGGGTAATAAT <i>G</i>
T507	CTACGAGGCAGATACAGTGATGNTACGTTATCAGACGACGCTTGTGGAGCACCTCACACATGTGAAGAGATGATCTAACAGCCTGGAAATAGCTGGGGGCGGTGGGGAGGA <u>R</u> GCTGGGCTTCC <i>C</i> CCTAGTAGCTCAGATGGTGAAGCATCTGCCTGCAATGCAAGAGACCCGGGTTCAATCCCTGGGTTGTGAAGATCCCTTGGAGAAGGAAATGGCAACCCACTCCAATATTTTGCCTGGAAAAATC CATGGATGGAAGAGCCTGGTAGGCTACAGTCCATGGGGTAGCAAAGAGTCACATATGACAGAGCAACTAACATATATACGTATACAGCTACGTAAATTTTATGTTCCCCAGTGAAGTCAAGCGTA <i>A</i> <i>A</i> AAGAATCTGCCT

Lokusz	Flanking régió szekvencia illesztései
C02	<p><u>CAGAGCTTGAGTGGTGGACTTAATGCCACAGCAGCCTCCAAGTGGGCGTCCCAACCC</u> <u>R</u> TGTATCTATCCATCTCCTCCTTGCAAGGTCCTTTGATGGATACTAGAAAAATAGAGAGCTCTCCTC A...-T.....CT..... TCTGCTAACTCCTAGTCTGGTGGAGATGTAGGTATGGAAATGAATTAGAAAAAAGTTCTTAAGCAAGGAATTAAGACACACTAAGATCTTACTATCAGCCAGGAGATCATGAGCCAAGAAATTAG C..... CTTTTATGGCTTTAGTGTACATACATTACAGCACATATCATATATTACATCTTGACTACATACATTCTTATAGCTAGTTTGGAGAGTTACACATAACAGTAGGTCACACATGTGAACCACCTTTGACG C.....A CATGTGACATGGAATAATCATTGCTGTGATGGAGGGAGACACAGAGGCT<u>TGTGAGGAACCCAGAGGAGAGGT</u> </p>
C143	<p><u>CAGGCGATCAAGAAATGTGGAATCAATGAAATCATCATTTTCCATTTTGCAAACAACTCAGGGNAAAGTAGAGGTATCTCTAAGGAGTCTTTCAGTTTTGAGAATATTTTACTGA</u> <u>R</u> ATGAATC C.....G...-..... CAATAGATGGAAGTTTAATAGGTT-CTCCATCCAGAAGCCACTGGTCCCAGACTTAGAAAGTAATGAAACAACAAGCAAAGACAGAACCTGCCTCCTGGTCCCCGACAAGAGCCATGGCTGTTGA T.....G.....-...G.....A..... ACTTGAGATAATTAATGCGTTTGCGAATTCCTGCTCCCTTCTCAGCCTGCCAGGTCTGCGAAGCCTCCTGCCAAGCACAGCAGGACAGATCAGCTCCACAGCCTCCAACACCT C.....G.....T.....</p>
C180	<p>GTGGGCATTCACTAGATCTTGG <u>R</u> GAGTAGGTGCTCAGTAATGTTTATGGAATGGACCCAAAGGTGTCTCTCCTAGACTCCATTATCTGAATTCCTTAAAGCAATGCCTTCTCTGCCTTATGA TTTAATGTTATGCTCTCTGTCTTCAGAGAACCCTGATGTGAGAACTGACCAACTTAATGGCATTCTGTCTTTATGAAAACAGAATTGAATATATTGCCCACTATGTCACTCAGTGTCTTTA C..... TACAAGGACTCATGATAAATTCATGCTGGGGGTTATTTCCAGGAAACACTTTGTCTCTGCAGATGAAGATGTTGTATTTTCTGCTGTGAGAAAATAAAAACTCTGGCTTCCGACGTTGTTTCCAT T.....-.....C.....-..... TTTTGCTTTAGCTGACAAATAGGGATACCCAGTTTTTATTACCTTATTACCTCCATGCATTCTGACC -.....</p>
T107	<p><u>ATTGACTGGTTTGATCTCCTTGCTAGAAAGGCTGAAGGAGAAGATTCATGTTTGAGATATATTTCTAAAGTATCATGTGAGGGCTCAGGGCTAAGATGTGATTTACGACTGTGTATGAGTCAAATC</u> A..... ATCAGAAGTTCCCTGGACACAGAATAGAGATATTGAAGATATGGACCAATTTGGAACATCCGTTTCAGGTGTGACAAGACACTGAAAAGTAGGATTGTCTTTTCCTCATAAAGAACAATTGGTCAG <u>R</u> G..... TATTGGTATATCTGAGCTTCCCCAAAAGGGCGTTATAGCTTCCTCTGAGGTATCTGCCAATAAGATGAGAAGATGTGGCTTTCTTCCTGGTCAAATCAGATCTGTTTACCATTATCTTACCTCT GGTCNAT<u>GTCTCCATATTAGTATCTCATTCATT</u> A..</p>

F.3. táblázat

O. hemionus és *C. canadensis* STR határoló régiók szekvencia összevetése a *C. elaphus hippelaphus* ortológgal. A GenBank-ban lévő REFERENCIA szekvencia; C...elaphus...hippelaphus szekvencia; a referencia szekvenciához viszonyított Mutációs helyek; bizonyos alléloknál detektált Mutációs helyek; a DeerPlex I and II PRIMER kötőhelyek; a mikroszatellita gének Repeat régiói; *N* ismeretlen nukleotid pozíciók. Klónozó PRIMER kötő helyek.

Lokusz	Allél	Repeat régió szerkezete	Amplikon
T108	16*	(GATA) ₁₆	<u>GTGGAGATAGGTAGACAGATGGAGAGAGGATAGA</u> GATATAT(GATA) ₁₆ GAGATAAATCAGGCAGAGAG AGCAAATACCATCTTTGGCAACCTCTGAATCAGC
	11	(GATA) ₁₁	
T172	15	(TATC) ₁₅	<u>CCCTTTCAACATGTCATTAGAGCCTAGAACAATT</u> (TATC) ₁₁ TATATGTATGTGGGCTTCC CAGGTAGCG CTAGTGGTAAAGAACA AAGAAAC
	11*	(TATC) ₁₁	
	10	(TATC) ₁₀	
	5	(TATC) ₅	
T123	19	(TAGA) ₁₄ N ₄ (TAGA)TGGG(TAGA) ₃	TTGTTCTGTTGCTTAGTTGTGTCTGACTCT(TAGA) ₁₅ TGGA(TAGA) ₃ CATAAATCATGTGGGCTCAGCTCAG
	18*	(TAGA) ₁₅ N ₄ (TAGA) ₃	
	16	(TAGA) ₁₃ N ₄ (TAGA) ₃	
	13	(TAGA) ₁₀ N ₄ (TAGA) ₃	
C229	16*	(CATC) ₉ (ATCC) ₇	TTATTCATCCACCCATCCATCACCATC(CATC) ₉ (ATCC) ₇ ATCATCCCTTCACTTATGAGCATGTGCC
	15	(CATC) ₈ (ATCC) ₇	
	13	(CATC) ₈ (ATCC) ₅	
T26	17	(TATC) ₁₄ TGTC(TATC) ₂ N ₂₇ (A) ₇ (GA)	CATGCAGATGCTACTCCCTTATGGGGAAAA(TATC) ₅ AGGTTAATTTCTGAATAATAGAGCCTAAAGGAA TAAGTGAAATAGTGTTATGTTTTG(A) ₇ GAATAAA GATTCAGGGTGCAGTATCTCTTCTGG
	15.1	(TATC) ₁₂ TGTC(TATC) ₂ N ₂₇ (A) ₆ (GA) ₂	
	13.1	(TATC) ₁₀ TGTC(TATC) ₂ N ₂₇ (A) ₆ (GA) ₂	
	13	(TATC) ₁₀ TGTC(TATC) ₂ N ₂₇ (A) ₇ (GA)	
	12.1	(TATC)TATT(TATC) ₇ TGTC(TATC) ₂ N ₂₇ (A) ₆ (GA) ₂	
	5*	(TATC) ₅ N ₂₇ (A) ₇ (GA)	
C01	27	AN ₂ (CATC) ₁₅ N ₅ (ATCC) ₄ N ₅₁ TN ₄ (ATCC) ₃ (AC) ₃ (CCAT) ₅	GTTATCCTTCTACCTTACATCCACCTGTT(CATC) ₁₄ CATTT(ATCC) ₄ ATTCATTTACCCACCTGCCTATCC ATTCATTCATCCATACATCCATCCACCCAGCT (ATCC) ₃ (AC) ₃ (CCAT) ₅ CTAACTATTCATCCATTTAT TTGAGATCCAGGTGTGGGAGG
	26*	GN ₂ (CATC) ₁₄ N ₅ (ATCC) ₄ N ₅₁ CN ₄ (ATCC) ₃ (AC) ₃ (CCAT) ₅	
	22	AN ₂ (CATC) ₁₀ N ₅ (ATCC) ₄ N ₅₁ TN ₄ (ATCC) ₃ (AC) ₃ (CCAT) ₅	
	20	AN ₂ (CATC) ₈ N ₅ (ATCC) ₄ N ₅₁ TN ₄ (ATCC) ₃ (AC) ₃ (CCAT) ₅	
	18	AN ₂ (CATC) ₆ N ₅ (ATCC) ₄ N ₅₁ TN ₄ (ATCC) ₃ (AC) ₃ (CCAT) ₅	

F.4-a. táblázat *Polimorf*, nem polimorf szekvencia motívumok; aláhúzott, primer szekvencia; aláhúzott, pig-tailed primer szekvencia; *, szekvenált allélok

Lokusz	Allél	Repeat régió szerkezete	Amplikon
T501	11.2*	<i>GGN₁₉(TCTA)TC(TCTA)₁₀N₆(TA)₃N₁₂(A)₁₀</i>	<u>CTCAGCCTACATGCTTTGACCAAGACCCG</u> GGGC AACCAATTAAATATATC(TCTA) TC(TCTA) ₁₀ CCTATC (TA) ₃ CACATACATAGT(A) ₁₀ TATGAGTCAGATTTGC ACAATATGCTTGTACA AAGAAAC
	10.2	<i>ATN₁₉(TCTA)TC(TCTA)₉N₆(TA)₃N₁₂(A)₁₀</i>	
	9.3	<i>ATN₁₉(TCTA)TC(TCTA)₈N₆(TA)₃N₁₂(A)₁₁</i>	
	6.1	<i>ATN₁₉(TCTA)₆N₆(TA)₃N₁₂(A)₁₁</i>	
	5.3	<i>ATN₁₉(TCTA)₆N₆(TA)₃N₁₂(A)₉</i>	
	5	<i>ATN₁₉(TCTA)₄(TCTG)N₆(TA)₃N₁₂(A)₁₀</i>	
T193	21	<i>(TGGA)₃(TAGA)₁₅(CAGA)(TAGA)₂</i>	<u>AGATGATAGACAGATGGGTAGCGGGTAGACAG</u> (TGGA) ₃ (TAGA) ₁₀ CAGACTGAGATTCCTCAGCT CTAAAATGAATGGTAATGATGACAACAGCAGT <u>AGAGC</u>
	18	<i>(TGGA)₄(TAGA)₁₄</i>	
	17	<i>(TGGA)₄(TAGA)₁₃</i>	
	13*	<i>(TGGA)₃(TAGA)₁₀</i>	
	11	<i>(TGGA)₃(TAGA)₈</i>	
	10	<i>(TGGA)₃(TAGA)₇</i>	
T507	14	<i>(GATA)₁₄N₂(TA)₅</i>	<u>GTTATCAGACGACGCTGTGGAGCACCTCACACATG</u> TGAAGAGATGATCTAACAGCCTGGAAATAGCTG GGGGCGGTGGGGAGGA(GATA) ₅ TG(TA) ₅ GCTGGG CTTCCCTAGTAGCTCAGATGGTGAAGCATCTGCC <u>TGCAATGCAAGAGACCCGGGTCAA</u>
	12	<i>(GATA)₁₂N₂(TA)₅</i>	
	10.1	<i>(GATA)₁₀AN₂(TA)₅</i>	
	10	<i>(GATA)₁₀N₂(TA)₅</i>	
	8	<i>(GATA)₈N₂(TA)₅</i>	
	5*	<i>(GATA)₅N₂(TA)₅</i>	
T156	14	<i>(GT)₂N₂₁TN₂₇(GATA)₁₄</i>	<u>CCTGGCCTGTGTCTTGAATTGAACCT</u> (GT) ₂ CTCCTG CATTGGCAGGCAGAT T CTTTACCACTGTGCCACC TGGGAAACT(GATA) ₆ TAGATCAGACAAGACTGGGT <u>ATTCATCGCC</u>
	13.2	<i>(GT)₁N₂₁CN₂₇(GATA)₁₄</i>	
	13	<i>(GT)₂N₂₁TN₂₇(GATA)₁₃</i>	
	12	<i>(GT)₂N₂₁TN₂₇(GATA)₁₂</i>	
	10	<i>(GT)₂N₂₁TN₂₇(GATA)₁₀</i>	
	9	<i>(GT)₂N₂₁TN₂₇(GATA)₉</i>	
	6*	<i>(GT)₂N₂₁TN₂₇(GATA)₆</i>	

F.4-b. táblázat

Polimorf, nem polimorf szekvencia motívumok; aláhúzott, primer szekvencia; aláhúzott, pig-tailed primer szekvencia; *, szekvenált allélok.

	T501					T172					T156					C01					T507				
Allél	Min	Max	d	SD	n	Min	Max	d	SD	n	Min	Max	d	SD	n	Min	Max	d	SD	n	Min	Max	d	SD	n
5	136.47	136.89	0.42	0,118	10	101.37	102.10	0.73	0.140	21	125.69	126.39	0.70	0.250	22						182.78	182.99	0.21	0.053	23
5.1	137.78	137.84	0.06	0,035	3																				
5.3	139.71	139.91	0.20	0,078	5																				
6	140.81	140.92	0.11	0,078	2						129.82	130.51	0.69	0.222	32						186.69	186.88	0.19	0.053	24
6.1	141.73	142.01	0.28	0,078	26																				
6.3	143.98	144.19	0.21	0,081	5																				
7	145.35	145.35	0.00	0,000	1	109.52	109.93	0.41	0.123	9											190.60	190.84	0.24	0.071	13
7.1	146.23	146.52	0.29	0,089	14																				
7.3	148.61	148.93	0.32	0,106	7																				
8	149.73	150.02	0.29	0,088	10	113.39	113.67	0.28	0.088	12	137.70	138.59	0.89	0.237	19						194.45	194.72	0.27	0.061	52
8.2	152.01	152.15	0.14	0,099	2																				
8.3	153.11	153.39	0.28	0,146	3																				
9	154.16	154.35	0.19	0,085	5	117.40	117.72	0.32	0.134	4	142.37	143.32	0.95	0.284	14						198.36	198.57	0.21	0.055	13
9.2	156.25	156.46	0.21	0,097	4																				
9.3	157.32	157.57	0.25	0,072	15																				
10						121.00	121.60	0.60	0.188	14	146.83	147.39	0.56	0.190	17						202.31	202.55	0.24	0.077	7
10.1	159.06	159.66	0.60	0,153	24																203.36	203.45	0.09	0.064	2
10.2	160.48	160.96	0.48	0,157	20																				
11	162.76	162.76	0.00	0,000	1	125.08	125.80	0.72	0.213	15	151.26	151.60	0.34	0.141	7										
11.2	164.73	165.04	0.31	0,106	30																				
11.3	165.72	165.94	0.22	0,105	5																				
12						129.10	130.01	0.91	0.252	14	155.46	156.03	0.57	0.163	22						210.06	210.35	0.29	0.079	23
12.2	169.12	169.12	0.00	0,000	1																				
13						133.74	134.09	0.35	0.098	17	160.05	160.05	0.00	0.000	1						214.09	214.22	0.13	0.053	7
13.1	171.90	172.04	0.14	0,081	3																				
13.2											161.86	162.19	0.33	0.104	16										
14						137.58	138.24	0.66	0.147	40	163.97	164.39	0.43	0.099	28						217.95	218.28	0.33	0.098	21
15						142.09	142.58	0.49	0.106	34	168.00	168.22	0.22	0.097	4						221.98	222.32	0.34	0.091	15
15.1	180.18	180.47	0.29	0,135	4																				
16						146.83	147.20	0.37	0.119	13	172.07	172.39	0.32	0.165	3										
17						151,32	151.45	0.13	0.041	6	176.10	176.35	0.25	0.069	11	202.64	203.05	0.41	0.114	12					
18						155,84	155,84	0.00	0.000	1	180.11	180.11	0.00	0.000	1	206.55	207.05	0.50	0.100	30					
19											184.07	184.29	0.22	0.090	4	210.44	210.96	0.52	0.130	20					
20																214.34	215.00	0.66	0.134	39					
21																218.29	218.71	0.42	0.134	13					
22																222.18	222.83	0.65	0.122	33					
23																226.18	226.63	0.45	0.165	6					
24																230.09	230.81	0.72	0.226	9					
25																234.01	234.24	0.23	0.100	4					
26																237.99	238.14	0.15	0.068	5					
27																241.86	242.45	0.59	0.171	28					
30																253.78	253.78	0.00	0.000	1					

	C229					T26					T193					T108					T123				
Allél	Min	Max	d	SD	n	Min	Max	d	SD	n	Min	Max	d	SD	n	Min	Max	d	SD	n	Min	Max	d	SD	n
5						144.32	144.46	0.14	0.048	12											79.60	79.60	0.00	0.000	1
6																115.30	115.71	0.41	0.148	7					
7						151.97	152.15	0.18	0.070	10						119.36	119.96	0.60	0.139	53					
8						157.66	157.91	0.25	0.092	6						123.43	123.90	0.47	0.110	36					
8.3						160.65	160.65	0.00	0.000	1															
9																127.27	128.02	0.75	0.122	43					
9.1						162.72	163.40	0.68	0.263	6															
10						165.74	165.87	0.13	0.054	4	127.01	127.23	0.22	0.081	11	131.45	131.87	0.42	0.105	13					
11											131.04	131.27	0.23	0.073	17	135.82	135.96	0.14	0.061	6	104.18	104.53	0.35	0.137	13
12						173.78	174.06	0.28	0.084	13						139.49	140.18	0.69	0.169	32	108.02	108.30	0.28	0.107	5
12.1						174.81	175.19	0.38	0.124	20															
13	100.05	100.73	0.68	0.144	33	177.77	178.11	0.34	0.089	34	139.00	139.60	0.60	0.120	28	144.58	144.61	0.03	0.021	2	111.61	112.16	0.55	0.133	43
13.1						178.95	178.95	0.00	0.000	2															
14	104.24	104.49	0.25	0.177	2	181.79	182.29	0.50	0.135	12	143.73	143.98	0.25	0.092	8	148.38	148.63	0.25	0.108	4	115.73	116.38	0.65	0.151	30
14.1						182.97	183.17	0.20	0.077	7															
15	107.78	108.53	0.75	0.147	135	185.65	185.83	0.18	0.081	4	148.20	148.45	0.25	0.074	13						119.44	120.22	0.78	0.158	19
15.1						186.69	187.25	0.56	0.123	23															
16	111.64	112.21	0.57	0.134	28	189.71	189.97	0.26	0.114	6	152.55	152.91	0.36	0.118	10	156.70	156.85	0.15	0.070	4	123.67	124.42	0.75	0.177	30
16.1						190.60	190.88	0.28	0.111	6															
17	115.94	115.97	0.03	0.021	2	193.53	193.95	0.42	0.119	26	156.77	157.30	0.53	0.173	16						127.47	128.18	0.71	0.196	10
17.1						194.76	194.76	0.00	0.000	2															
18						197.58	197.58	0.00	0.000	1	161.09	161.51	0.42	0.127	21						131.68	132.42	0.74	0.175	12
18.1						198.49	198.80	0.31	0.143	4															
19						201.48	201.48	0.00	0.000	1	164.28	164.28	0.00	0.000	1						136.03	136.66	0.63	0.145	37
20											169.11	169.44	0.33	0.099	27										
21											173.17	173.40	0.23	0.075	18										
22											177.13	177.43	0.33	0.114	6										
23											181.05	181.41	0.36	0.109	17										
24											185.12	185.29	0.17	0.083	5										
26											193.03	193.13	0.10	0.092	2										

F.5-a., F.5-b. táblázat *DeerPlex II* (F.5-a. 89. oldal) és *-I* (F.5-b. 90. oldal) allélikus fragmens csoportok. *Min* és *Max*, minimum és a maximum megfigyelt fragmens méret; *d*, *Min* és a *Max* értékek közötti különbség; *SD*, szórás; *n*, a detektált allélszám.

Minta	T108		T172		T123		C229		T26		C01		T501		T193		T507		T156	
EK01	7	8	8	17	12	15	16	17	8	15	20	24	7.3	10.2	10	20	12	13	13.2	15
EK02	7	9	13	15	18	19	14	16	15.1	16.1	19	20	5	8	14	15	12	12	6	10
EK03	9	9	15	16	16	19	16	16	12.1	13	19	27	8.2	10.1	14	16	10.1	12	5	12
EK04	8	12	13	15	11	15	16	17	9	18.1	21	27	9	11.2	20	21	5	12	13.2	14
EK05	7	8	5	16	16	19	14	16	15.1	16.1	18	20	10.1	11.3	20	20	5	15	14	14
EK06	8	9	8	15	13	13	14	17	5	13	17	18	7.1	11.2	10	20	8	15	14	14
EK07	8	13	5	5	17	18	16	16	10	12.1	18	23	7	9	19	20	5	7	6	8
EK08	7	7	5	12	15	17	14	16	5	13	27	27	8	12.2	16	17	5	6	6	12
EK09	7	8	15	15	11	16	14	16	12	17	22	23	7.3	9.3	18	20	6	6	8	13.2
EK10	6	8	7	15	19	19	14	16	12	14	20	23	11	11.3	18	21	6	13	6	6
EK11	9	10	5	16	14	16	16	16	10	12.1	21	27	7.1	15.3	13	17	5	7	14	14
EK12	7	8	15	16	13	19	16	16	8	15.1	18	23	5	6	14	17	6	7	6	12
EK13	7	7	7	17	14	16	17	17	12	15.1	20	27	9.2	13.3	13	20	5	8	14	14
EK14	8	10	8	17	15	16	16	16	12	15.1	24	27	5	8	16	18	5	5	13.2	14
EK15	7	8	5	16	13	17	16	16	13	17	19	27	6.1	13.3	11	13	7	14	9	17
EK16	7	14	8	14	16	19	14	16	9	15.1	19	20	7.3	10.2	17	23	6	7	8	13.2
EK17	8	9	5	13	17	19	16	17	13	17	27	27	8	8.3	17	24	7	10	12	14
EK18	7	8	14	15	14	16	17	17	12.1	15	22	24	11.3	15.3	15	20	5	8	6	8
EK19	8	8	14	17	18	19	14	16	8	12.1	27	27	6.1	9.2	10	17	5	8	6	12
EK20	7	10	16	16	13	16	16	16	12.1	17	20	23	5	10.1	10	21	6	14	5	6
EK21	7	8	5	14	14	16	16	17	9	18.1	19	20	6.3	10.1	10	10	5	8	6	14
EK22	7	10	5	7	16	19	17	18	14	16.1	20	27	6.1	10.2	18	20	7	15	6	14
EK23	7	12	15	16	15	16	16	16	9	13	19	19	10.2	10.2	18	20	7	13	9	9
EK24	8	8	7	14	16	19	16	16	8.2	12.1	19	27	6.1	8.3	10	21	5	6	6	14
EK25	7	9	12	15	11	17	16	17	12	14.1	20	21	6.1	10.2	10	13	8	10	6	11
EK26	7	8	11	14	14	19	16	16	9	17	22	27	10.2	10.2	15	17	13	15	5	14
EK27	7	7	15	15	5	13	16	16	14	18.1	17	27	5	9	13	16	8	13	6	9

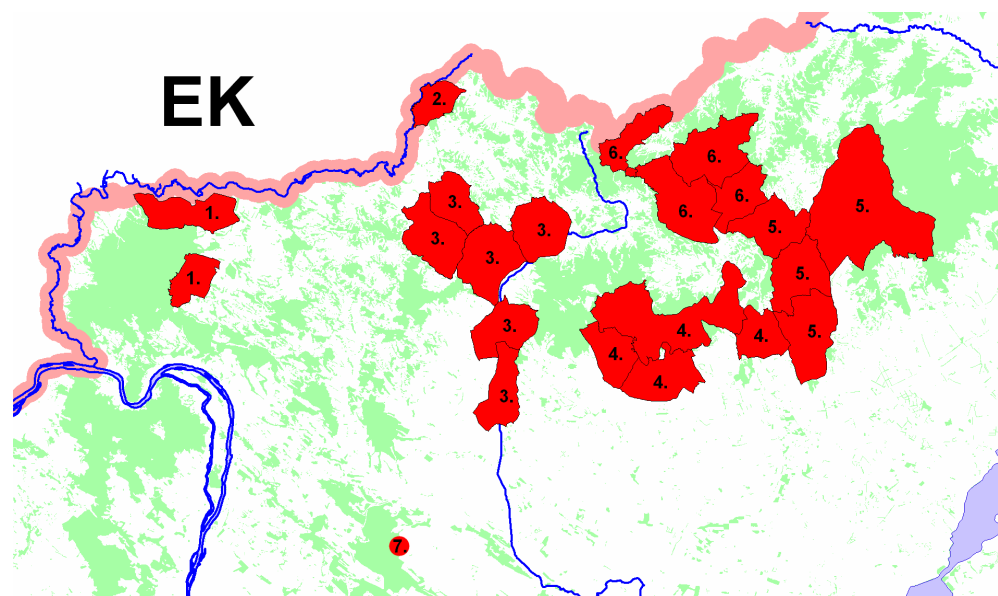
Minta	T108		T172		T123		C229		T26		C01		T501		T193		T507		T156	
EK28	7	7	5	12	11	15	14	14	12.1	16.1	21	27	7.3	10.2	20	23	10	12	10	16
EK29	8	8	12	13	16	19	16	16	14	17	19	27	7.3	10.2	16	20	7	10	6	14
EK30	10	13	5	15	11	16	14	15	18.1	19	21	21	5	6.1	14	17	6	8	8	10
EK31	8	8	15	15	13	18	14	16	12.1	12.1	18	24	7.1	11.2	13	20	7	15	14	17
EK32	7	12	15	17	16	17	16	16	14	15.1	27	27	8	8	16	21	5	13	10	13.2
EK33	8	8	15	15	16	16	14	16	5	17	17	18	11.2	11.2	10	13	8	15	6	14
EK34	9	11	15	17	13	16	16	16	15.1	16.1	21	27	5.1	15.3	16	20	5	8	6	14
EK35	6	7	13	14	17	19	16	16	8	12	17	18	5	10.1	22	23	6	10	12	13.2
EK36	7	14	5	14	14	16	14	17	13	14	19	24	7.1	8	16	18	10	15	10	12
EK37	7	14	5	5	14	16	14	16	13	13	22	27	7.3	10.1	15	20	15	15	13.2	17
EK38	7	8	5	16	16	19	14	16	15.1	16.1	18	20	10.1	11.3	20	20	5	15	14	14
EK39	7	8	13	16	14	14	15	16	12	12.1	20	24	8	15.3	18	18	12	12	6	13.2
EK40	7	14	9	15	14	17	17	17	12.1	15.1	19	27	6.3	7.1	20	24	5	10	13.2	14
EK41	7	12	9	13	15	15	16	16	12.1	12.1	17	21	6.3	10.1	14	16	10.1	12	5	6
EK42	9	12	10	11	15	19	16	16	7	10	17	22	9.3	11.2	13	15	6	8	5	10
EK43	9	11	8	14	14	14	16	17	9	12	18	18	9.3	11.2	13	15	8	12	6	8
EK44	7	7	12	14	13	18	16	17	14.1	15.1	18	26	9.3	11.2	11	13	8	8	8	12
EK45	8	12	7	13	14	19	16	16	7	12	18	22	10.1	11.2	13	21	8	12	12	17
EK46	6	9	5	5	16	19	16	16	12.1	14.1	19	27	7.1	10.1	10	17	5	12	13.2	15
EK47	9	9	14	15	13	15	16	16	5	15.1	18	18	7.1	9.2	14	24	8	9	8	13.2
EK48	8	9	10	14	14	16	14	16	15.1	15.1	18	22	7.1	10.1	13	15	8	8	10	13.2
DNY01	9	12	7	14	13	15	16	17	5	5	20	22	10.1	11.2	13	18	14	14	18	19
DNY02	8	12	15	16	13	15	16	16	14	17	18	21	5	9.2	13	13	5	8	6	17
DNY03	11	16	8	11	13	18	16	17	5	13	19	26	6.1	11.2	17	21	12	15	5	14
DNY04	12	12	7	12	11	16	14	16	7	13	19	24	5.1	11.2	11	13	14	14	12	16
DNY05	9	9	8	14	12	14	16	17	7	15.1	19	22	6	9	11	17	12	14	6	14
DNY06	12	12	5	14	13	15	16	16	14	15.1	18	20	9.3	11.2	18	20	6	14	12	13.2

Minta	T108		T172		T123		C229		T26		C01		T501		T193		T507		T156	
DNY07	9	11	8	16	14	19	16	16	13	17	17	22	5.1	5.3	20	22	8	15	6	17
DNY08	12	16	11	14	13	15	14	16	7	18	20	23	6.3	10.1	11	11	8	9	8	15
DNY09	7	12	10	14	13	13	16	17	15.1	16	22	25	7.1	10.1	18	26	8	14	6	19
DNY10	9	10	11	14	19	19	16	16	16	16	22	25	6.1	11.2	15	23	6	8	5	8
DNY11	7	12	9	10	13	14	14	14	7	13	21	22	10.2	11.2	20	23	8	8	8	12
DNY12	9	16	5	14	13	18	16	16	14	17.1	22	25	9.3	10.1	23	24	6	15	8	10
DNY13	9	10	12	12	15	19	16	17	13	17.1	17	21	11.2	11.2	13	16	8	8	10	19
DNY14	7	12	7	14	12	13	14	16	5	5	19	26	6.1	11.2	11	11	8	9	9	10
DNY15	9	12	5	12	13	14	16	16	15.1	17	21	22	6.1	6.3	23	23	7	14	5	12
DNY16	7	12	14	14	11	13	16	18	14.1	17	22	26	6.1	7.3	20	22	6	8	8	17
DNY17	7	9	13	14	13	16	14	16	17	17	18	20	11.2	11.2	17	21	7	14	8	17
DNY18	9	12	10	10	13	18	16	17	7	13	22	22	6.1	11.2	18	22	9	14	11	14
DNY19	7	12	12	13	13	13	14	16	5	13	18	27	10.1	11.2	20	22	8	12	12	19
DNY20	7	7	11	14	14	14	14	16	13	17	20	20	6.1	6.1	11	18	8	9	5	12
DNY21	12	12	14	15	11	14	16	16	13	17	18	21	10.2	10.2	21	23	8	14	5	6
DNY22	6	7	8	14	18	19	16	17	13	16	20	20	6.1	9.3	13	21	5	9	5	12
DNY23	7	9	14	14	13	19	14	16	13	13	20	20	10.1	10.1	13	15	9	14	6	6
DNY24	7	12	11	15	13	13	16	16	8	17	20	27	7.1	10.2	11	18	8	8	5	6
DNY25	9	10	10	11	11	15	16	17	13	15.1	20	20	6.1	6.1	11	21	6	8	5	9
DNY26	7	12	10	14	13	19	16	16	15	15.1	17	22	10.1	10.1	18	20	6	6	5	17
DNY27	7	9	11	14	14	15	14	16	15.1	17	20	22	6.1	6.1	11	18	8	13	12	12
DNY28	9	10	12	13	14	18	16	16	12.1	13	20	22	6.1	7.1	11	24	6	9	5	5
DNY29	10	10	11	14	13	14	17	17	8	12.1	20	22	9.3	10.2	11	18	5	8	5	9
DNY30	7	9	10	11	13	15	16	17	13	15.1	20	20	11.2	11.2	13	15	9	12	5	9
DNY31	9	12	13	14	11	15	16	16	16	17	20	22	10.2	10.2	10	13	6	12	5	9
DNY32	10	11	13	15	13	19	16	16	17	17	18	22	9	11.2	13	21	9	14	13.2	14
DNY33	6	9	13	15	14	18	16	16	12	13	18	18	9.3	11.2	11	23	8	8	11	12

Minta	T108		T172		T123		C229		T26		C01		T501		T193		T507		T156	
DNY34	9	9	10	13	19	19	16	16	12	13	22	22	6.1	6.1	15	18	8	12	5	9
DNY35	9	11	12	14	14	19	14	16	12	14	17	20	7.1	9.3	17	20	8	9	9	17
DNY36	9	9	8	18	13	19	16	16	13	15	19	22	8.2	9.3	18	23	12	15	9	13.2
DNY37	8	12	7	15	11	19	16	16	12	13	20	22	6.1	10.2	11	21	8	12	5	5
DNY38	9	12	14	15	11	19	16	16	13	15.1	18	20	9.3	11.2	13	23	8	8	6	6
DNY39	6	12	10	13	13	13	14	16	13	14.1	20	30	5.3	5.3	15	21	8	12	8	12
DNY40	12	12	10	15	19	19	16	16	13	14.1	17	18	9.3	9.3	13	23	6	8	10	14
DNY41	7	12	10	14	13	13	14	16	13	17	22	26	11.2	11.2	14	22	9	14	6	9
DNY42	8	8	12	14	12	13	16	17	13	17	22	24	10.1	10.1	11	23	8	14	11	11
DNY43	9	9	14	15	14	19	14	16	13	17	20	25	10.1	11.3	21	21	8	8	10	10
DNY44	7	7	14	15	14	18	16	16	12.1	17	17	27	8.3	13.3	18	18	8	14	8	10
DNY45	7	8	5	9	14	16	14	16	14	14.1	20	22	6.1	10.1	13	20	5	15	6	14
DNY46	7	8	8	15	13	13	14	16	12.1	12.1	18	20	8	10.2	17	26	6	7	10	16
DNY47	9	12	11	11	13	16	16	16	10	13	19	24	5	7.1	15	21	6	14	12	15
DNY48	9	9	8	13	13	19	16	16	14	17	18	22	6.1	11.2	23	23	12	14	8	11
DNY49	6	16	11	16	12	14	16	16	5	7	18	18	5.3	9.3	13	14	5	8	8	8
DNY50	7	10	10	11	16	16	16	16	16	17	18	22	5	8	13	13	5	6	12	13
DNY51	7	9	12	14	17	19	16	16	13.1	13.1	19	20	6.1	7.1	21	23	8	12	10	17
DNY52	8	8	14	15	11	17	16	16	7	7	19	27	5.3	10.1	17	17	9	14	9	10

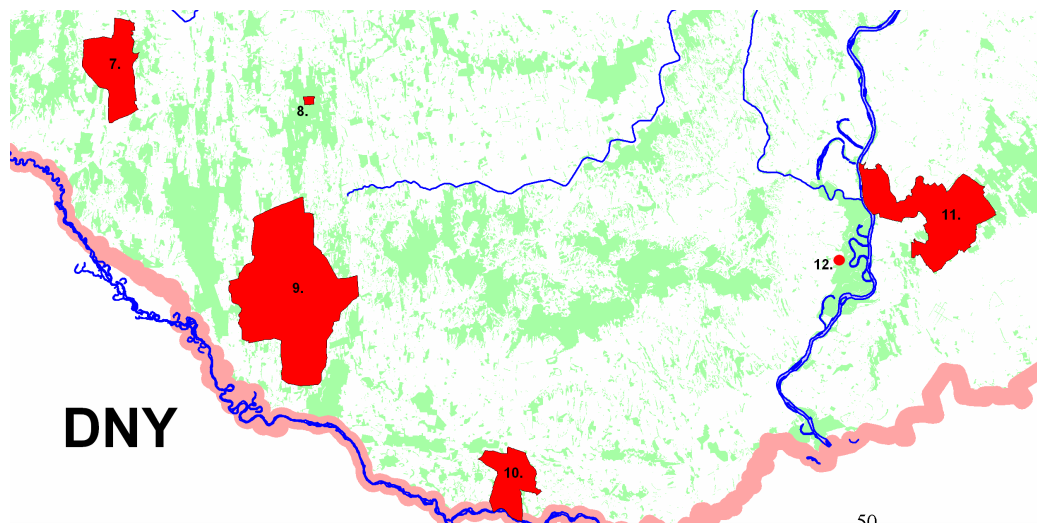
F.6-a., F.6-b., F.6-c., F.6-d. táblázatok *DeerPlex I-II* genetikai profilok: **F.6-a.:** EK01-EK27, **F.6-b.:** EK28-DNY07, **F.6-c.:** DNY08-DNY35, **F.6-d.:** DNY36-DNY52.

F.1. ábra Az északkelet-magyarországi (EK) minták származási helyei és típusai

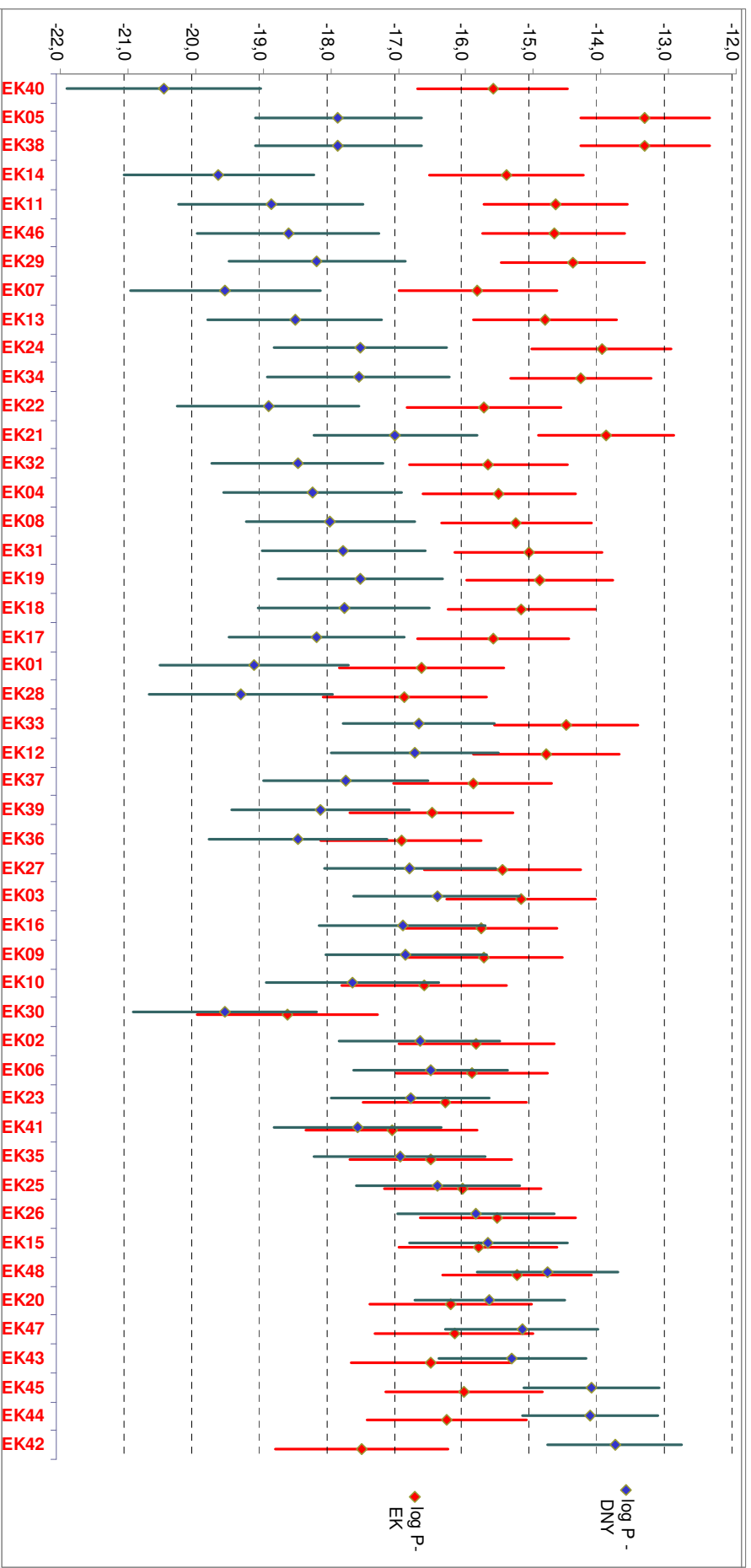


Terület	Mintaszáma	Mintakód	Mintatípus		
			Vér	Izom	Agancs
1.	3	EK09,EK26,EK27	x		
2.	3	EK19,EK35,EK37	x		
3.	11	EK04,EK05,EK06,EK08,EK21,EK31,EK32,EK33,EK34,EK36,EK38	x		
4.	6	EK13,EK15,EK17,EK18,EK40,EK41	x		
5.	9	EK01,EK02,EK03,EK11,EK12,EK14,EK20,EK22,EK24	x		
6.	9	EK07,EK10,EK16,EK23,EK25,EK28,EK29,EK30,EK39	x		
7.	7	EK42,EK43,EK44,EK45,EK46,EK47,EK48	x		

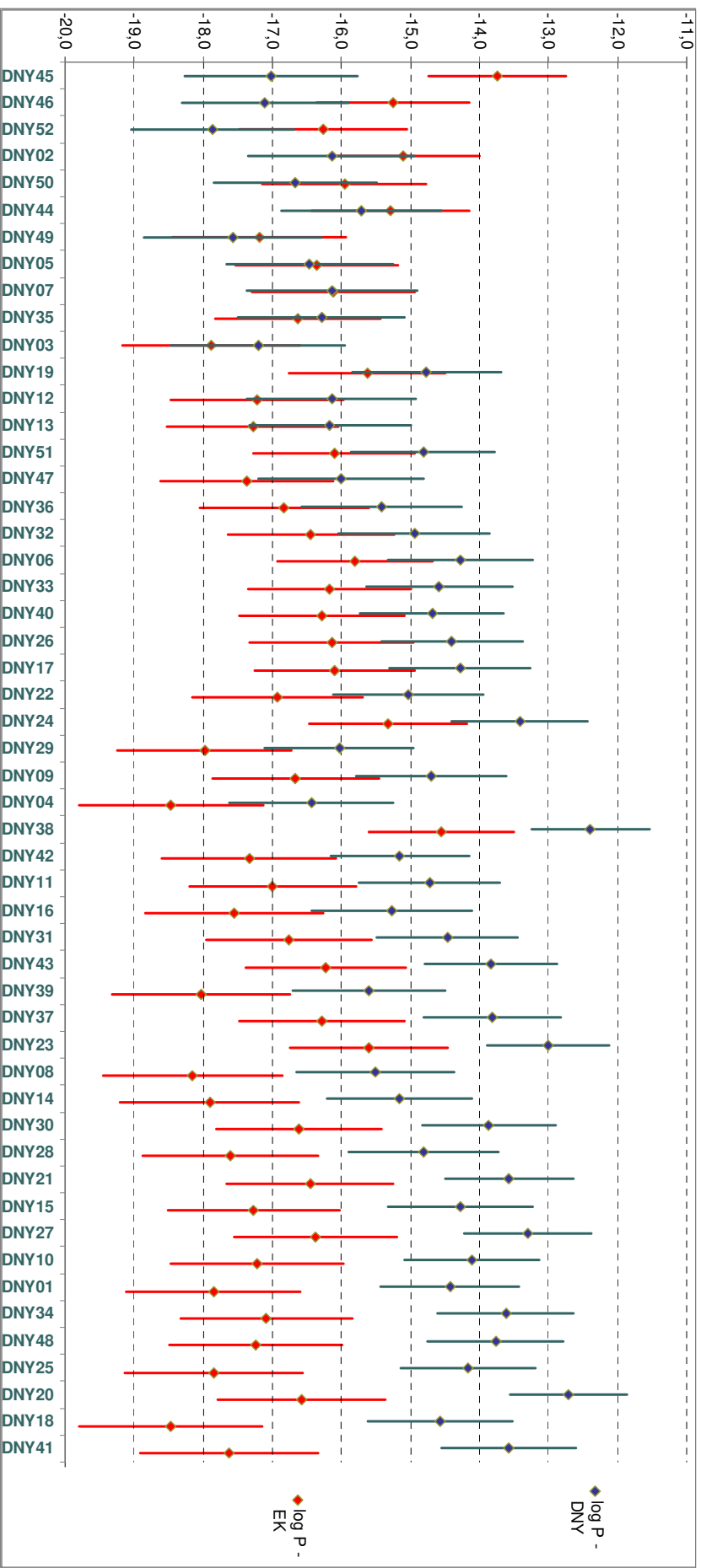
F.2. ábra A délnyugat-magyarországi (DNY) minták származási helyei és típusai



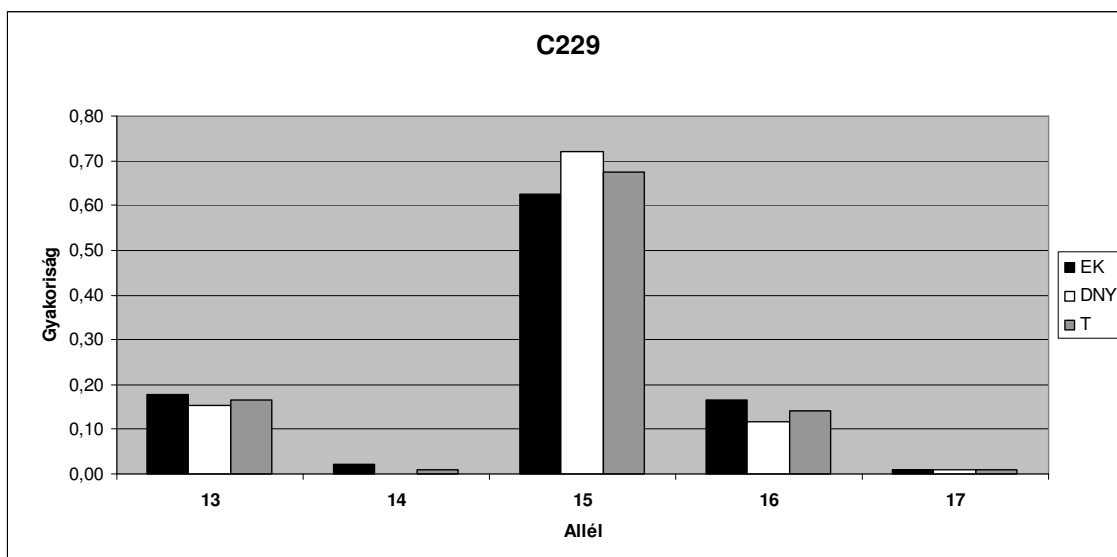
Terület	Mintaszám	Mintakód	Mintatípus		
			Vér	Izom	Agancs
7.	7	DNY04,DNY47,DNY48,DNY49,DNY50,DNY51,DNY52	1		6
8.	1	DNY03		1	
9.	17	DNY01,DNY05,DNY06,DNY07,DNY08,DNY09,DNY10,DNY11,DNY12 ,DNY13,DNY14,DNY15,DNY16,DNY17,DNY18,DNY19,DNY32		17	
10.	10	DNY33,DNY34,DNY35,DNY36,DNY37,DNY38,DNY39,DNY40,DNY41,DNY42	10		
11.	12	DNY20,DNY21,DNY22,DNY23,DNY24,DNY25, DNY26,DNY27,DNY28,DNY29,DNY30,DNY31		12	
9.,10.,11.	4	DNY43,DNY44,DNY45,DNY46	4		
12.	1	DNY02			1



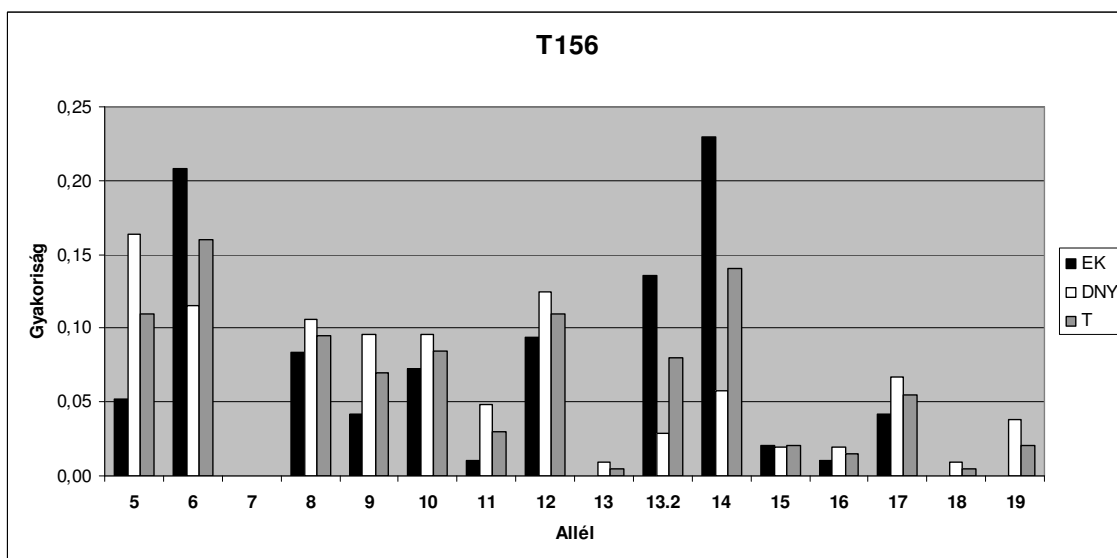
F.3-a. ábra Az északkeleti populációs minták EK és DNY populáció alapján kalkulált torzítatlan profilgyakorisági (P) értékeinek 10-es alapú logaritmusa és 95 %-os konfidencia intervalluma.



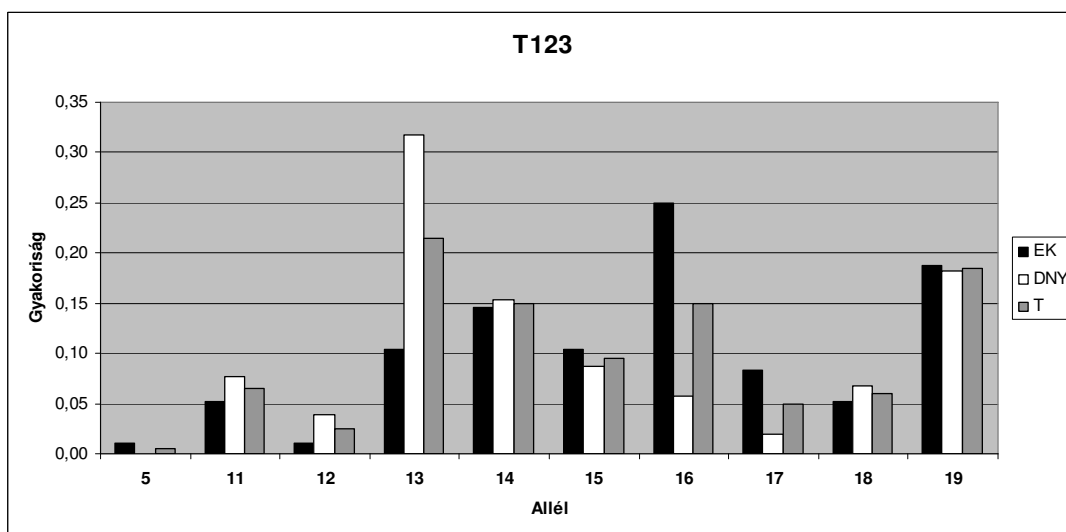
F.3-b. ábra A délnyugati populációs minták EK és DNY populáció alapján kalkulált torzítatlan profilgyakorisági (P) értékeinek 10-es alapú logaritmusá és 95 %-os konfidencia intervalluma.



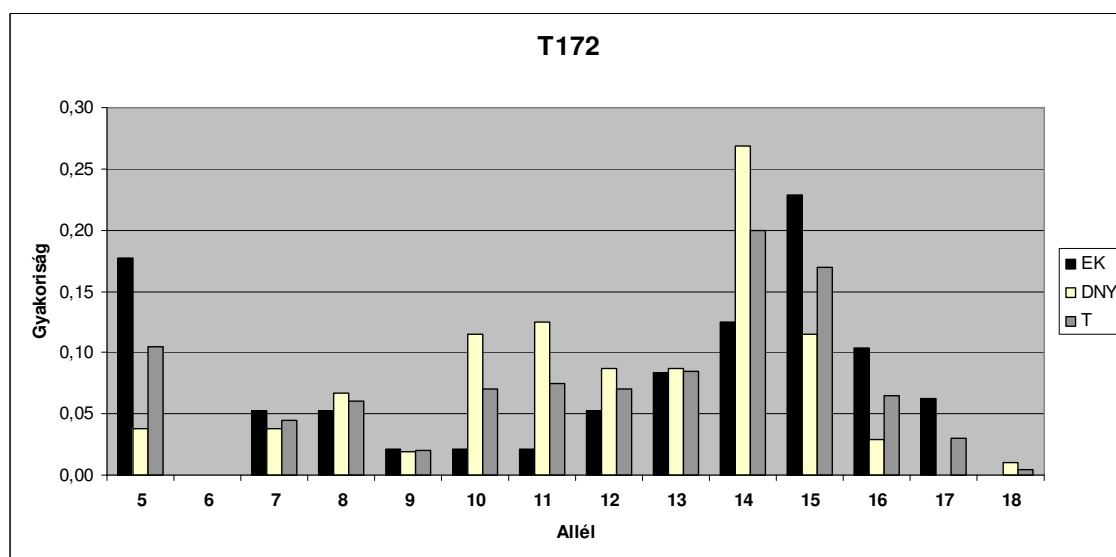
F.4. ábra C229 populációs allélgyakoriságainak eloszlása.



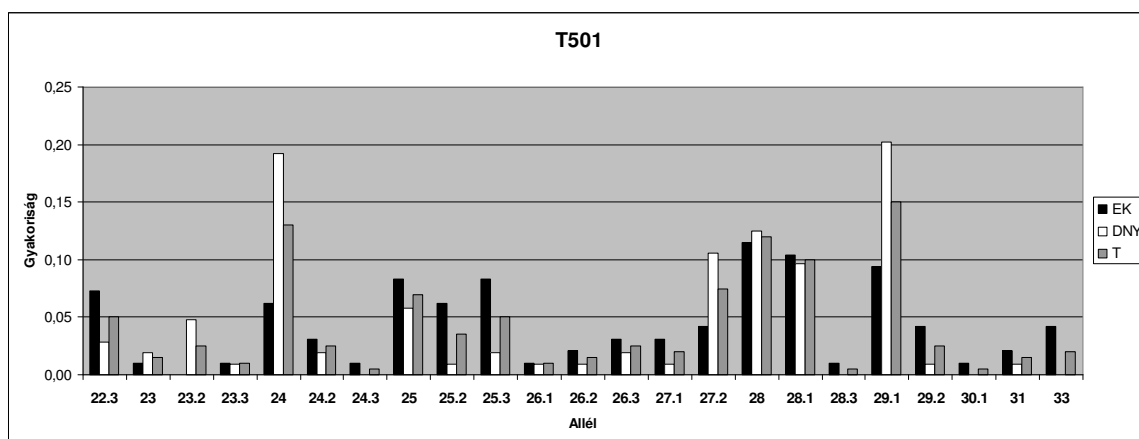
F.5. ábra T156 populációs allélgyakoriságainak eloszlása.



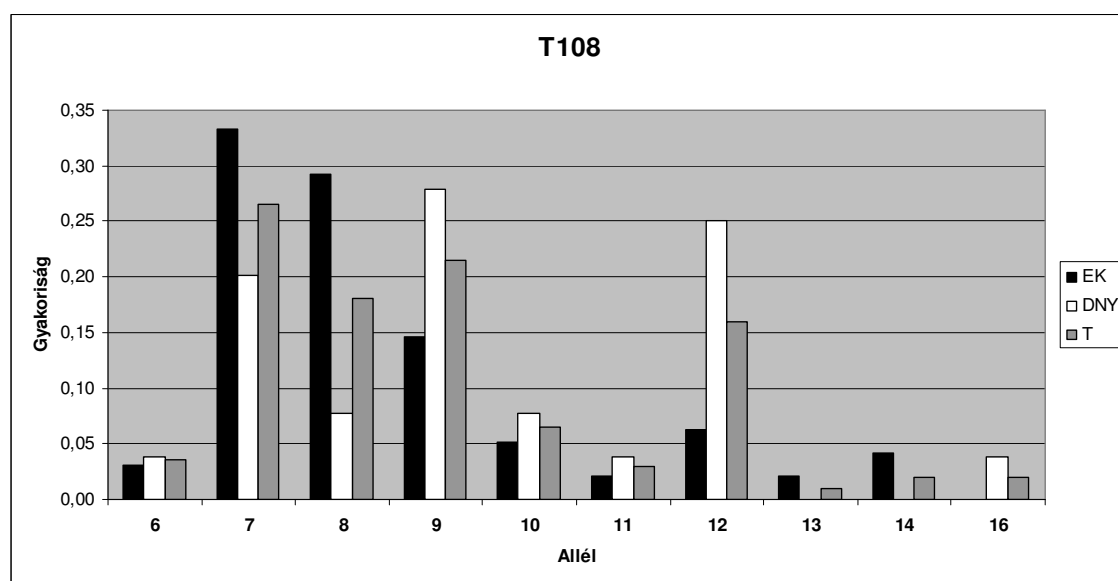
F.6. ábra T123 populációs allélgyakoriságainak eloszlása.



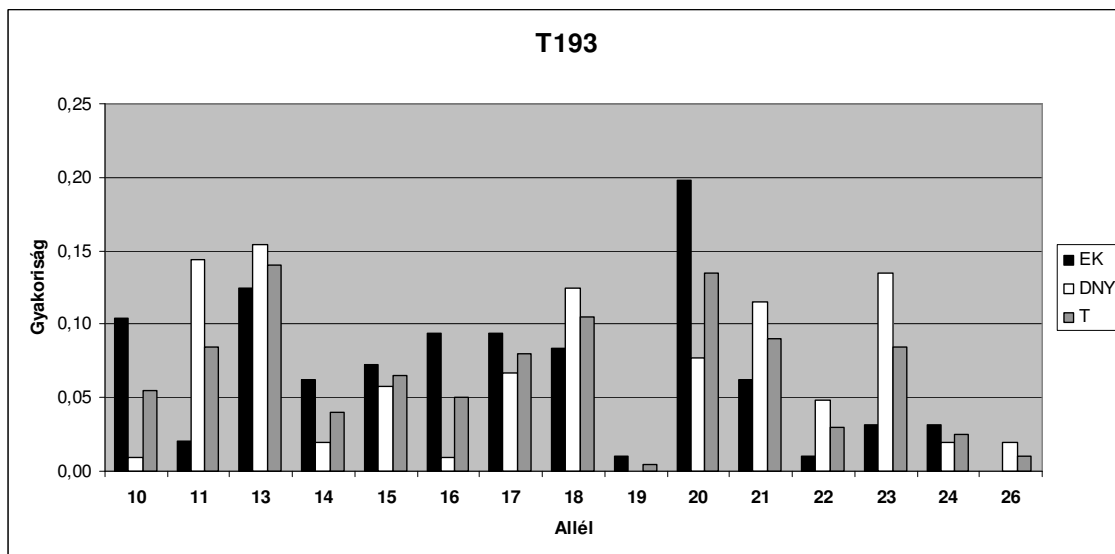
F.7. ábra T172 populációs allélgyakoriságainak eloszlása.



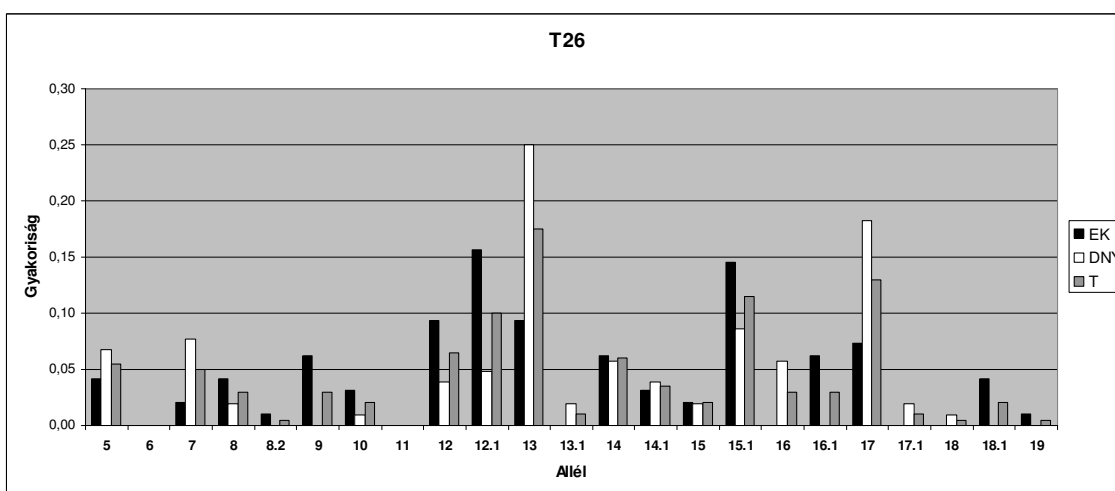
F.8. ábra T501 populációs allélgyakoriságainak eloszlása.



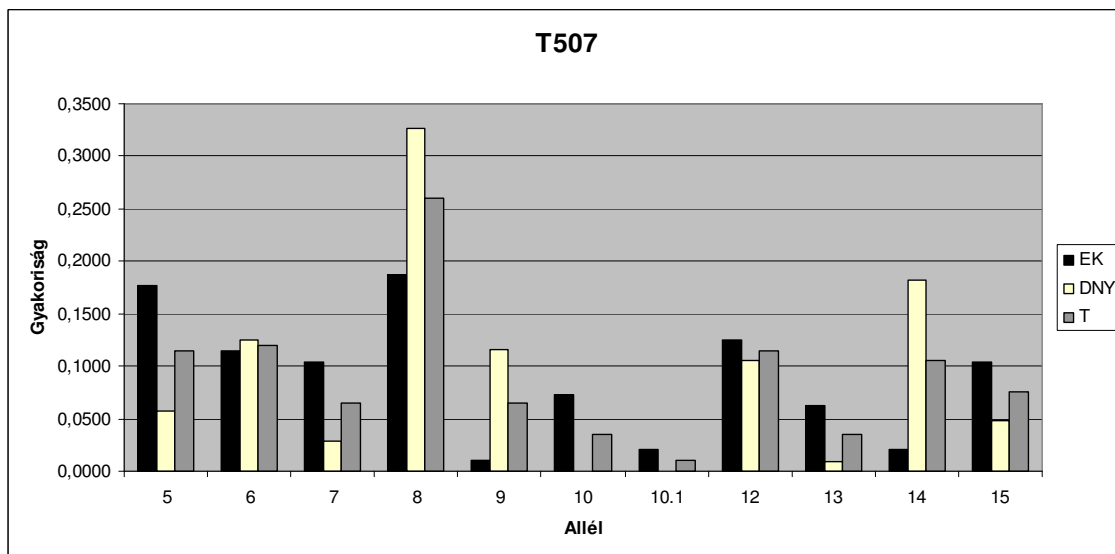
F.9. ábra T108 populációs allélgyakoriságainak eloszlása.



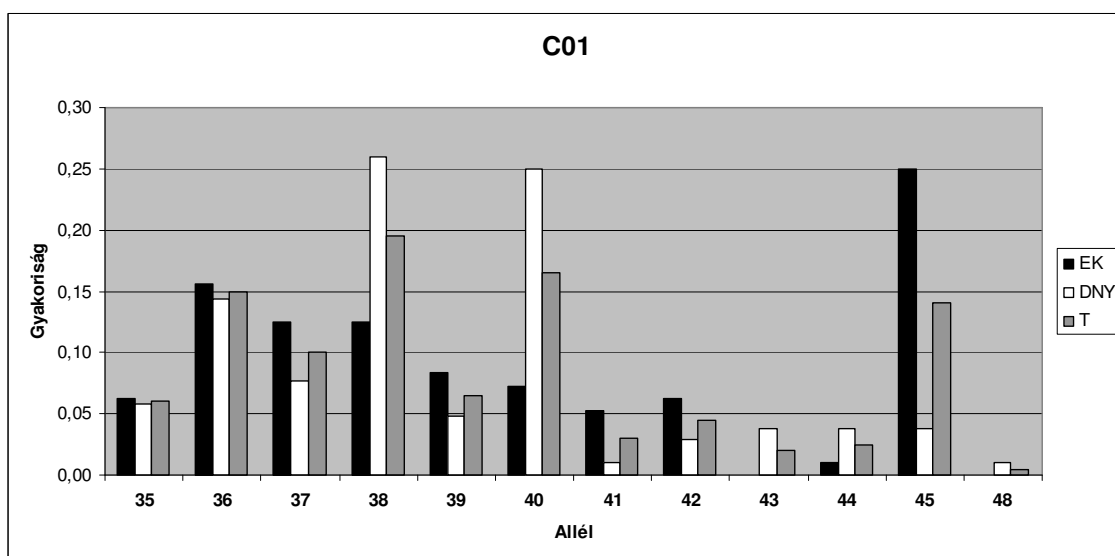
F.10. ábra T193 populációs allélgyakoriságainak eloszlása.



F.11. ábra T26 populációs allélgyakoriságainak eloszlása.



F.12. ábra T507 populációs allélgyakoriságainak eloszlása.



F.13. ábra C01 populációs allélgyakoriságainak eloszlása.

8 Összefoglalás

A gímszarvas (*Cervus elaphus hippelaphus*), mint nagytetű herbivor hazai faunánk ökológiai és vadgazdálkodási szempontból is nagy jelentőséggel bíró tagja. Becsült állománya 100.000 feletti egyedszámmra tehető, amelyből évente legálisan jelentős mennyiséget, hozzávetőlegesen 40.000 példányt ejtenek el. Az illegális veszteség a szakemberek becslései szerint igen magas lehet, évente akár 10.000 gímszarvas is orvvadászok zsákmánya lehet. Egyik célunk az volt, hogy egy olyan - gímszarvasok egyedazonosítására alkalmas – tetranukleotid (tetramer) mikroszatellitákból (STR) álló multiplex PCR rendszert hozzunk létre, amellyel az elkövetőt a tettenérés szüksége nélkül lehessen összefüggésbe hozni közvetlenül a tett helyszínével. A mikroszatellita lokuszoknak megfelelő polimorfizmus fokkal kell rendelkezniük ahhoz, hogy egy biológiai nyomot egy egyedtől származtatni tudjunk.

A lokuszok polimorfizmusának felméréséhez a délnyugati megyékből (DNY - Zala, Somogy, Baranya és Bács-kiskun (Gemenc)) valamint az ország északi területeiről (EK - Börzsöny, gödöllői-dombság, Cserhát, Bükk, Mátra) 100 mintát gyűjtöttünk. A munkánk másik célja a magyarországi gímszarvas állomány, mint metapopuláció genetikai struktúrájának részletesebb megismerése volt.

A szakirodalomban rendelkezésünkre álló amerikai vapitiből (*Cervus canadensis*) és őszvérszarvasból (*Odocoileus hemionus*) izolált 22 tetramer STR ortológból az ún. „zoo-klónozási” eljárással 14 lokusz szekvenciáját sikerült meghatározni. A PhastSystem és a nagy-felbontású PAGE technológia segítségével végzett polimorfizmus szűréssel 4 STR markert szelektáltunk ki. A 10 polimorf mikroszatellita lokusból két 5 STR markert tartalmazó multiplex PCR rendszert készítettünk (*DeerPlex I-II*). A markerek jelöléséhez a DS-33 (Applied Biosystems) standard mátrix fluorofórijait választottuk. Genotipizáló rendszernek az ABI 3130 kapilláris-gélelektroforézis készüléket és a Genemapper® ID 3.2 szoftvert (Applied Biosystems) alkalmaztuk.

A markerek identifikáló erejének becsléséhez megvizsgáltuk lokuszaink polimorfizmus fokát, amelyet a Nei-féle géndiverzitással, polimorfizmus információtartalommal és a megkülönböztető erővel jellemeztünk. A C229 lokuszt kivéve egyenként magas értékeket kaptunk. A lokuszokon belül az allélok random asszociációját a Hardy-

Weinberg egyensúly (HWE) tesztel az Arlequin v3.11 és a Genepop v4.0 szoftver felhasználásával végeztük. Az 5 %-os szignifikancia szinten a T501, C01 és T108 markereknél volt elvethető a HWE. A lokuszok függetlenségi vizsgálatait sem statisztikai (LD teszt) sem szekvencia hasonlóságon alapuló térképezés alapján nem zárták ki a C01 és a T26 genetikai kapcsoltságát, ezért az igazságügyi számításokhoz a C01 markert nem tartalmazó 9 lokuszt használtuk.

DeerPlex I-II egyedazonosító képességének vizsgálatához szükség volt a populációk genetikai struktúrájának ismeretére. Az Arlequin szoftver segítségével molekuláris variancia analízist végeztünk, amelynek eredményeképpen a két alpopuláció között szignifikáns genetikai elkülönülést mutattunk ki ($F_{ST} = 0.034$), az alpopulációkon belüli átlagos teljes beltenyészet paraméterre, $F_{IS} = 0.037$ értéket kaptunk. A két paraméterrel korrigált átlagos egyezési/azonosság valószínűséggel (pM/PI_{ave}) valamint a teoretikusan leggyakoribb genetikai profil egyediségével („uniqueness”) fejeztük ki a *DeerPlex I-II* identifikáló erejét. Minimum $1 - 7,93 \times 10^{-4}$ annak a valószínűsége, hogy egy 9 lokuszos *DeerPlex I-II* profil egyedi és $1,0056 \times 10^{-12}$ a valószínűsége annak, hogy a populációból két véletlenszerűen kiválasztott egyed 9 lokuszos genotípusa megegyezik. A vizsgálatok eredményeképpen igen magas identifikáló erőt kaptunk, megközelítőleg olyan erős, mint az Egyesült Államok kaukázusi népességére számolt Identifiler igazságügyben használt kit egyedazonosító ereje.

A marker készlet igazságügyi alkalmazhatóságát jól példázza az általunk megoldott orvvadászati eset is, amelyben a helyszíni nyomokból kimutatott *DeerPlex I-II* genetikai profilokat és a gyanúsított birtokán lefoglalt gímszarvasbika trófeáját megfelelő statisztikai erővel tudtuk egy egyedtől származtatni.

A *DeerPlex I-II* rendszereinket felhasználtuk az alpopulációk genetikai struktúrájának vizsgálatára is, amit az egyedek populációs eredetének valószínűsítésével – az ún. „assignment tesztel” - végeztünk el. Az EK-DNY adatbázisból kalkulált profilgyakorisági értékpárok alapján mindkét populációt alkotó egyedek 17 %-a nem a referencia populációba sorolt, ebből következően heterogénnek tekinthetők. Igazságügyben DNS alapú bizonyítékok statisztikai interpretációjához az alpopulációk - genetika struktúráltságuk következtében - nem használhatók, helyettük a teljes adatbázist ajánljuk. Az „assignment teszt” alapú populációs eredet meghatározás az egyedek két populáció szerint kalkulált profilgyakorisági értékeinek nagyarányú (44%)

átfedő konfidencia intervalluma alapján torzított lenne, ezért igazságügyi statisztikai interpretációhoz nem alkalmazható.

9 Summary

Being a large herbivorous species, the red deer (*Cervus elaphus hippelaphus*) is an important member of the Hungarian fauna from the point of view of ecology as well as of game management. The number of individuals is estimated to be more than 100,000. Approximately 40,000 of them are legally hunted each year. Illegal loss by poaching may be as high as 10,000. One of our objectives was to develop a multiple PCR system consisting of tetranucleotide (tetramer) microsatellites (STR) suitable for linking the offender directly with the crime scene without the necessity of catching the perpetrator in the act. To facilitate the linking of a biological trace with an individual, microsatellite loci must be sufficiently polymorphic.

To assess the polymorphisms of loci, a total of 100 samples were collected from the south-western counties (DNY – Zala, Somogy, Baranya and Bács-Kiskun (Gemenc)) and the north-eastern regions (EK – Börzsöny, Gödöllő Hills, Cserhát, Bükk and Mátra) of Hungary. The second objective of our investigation was gaining a more extensive knowledge of the genetical structure of the Hungarian red deer stock as a metapopulation.

Using the so called "zoo cloning" method, 14 locus sequences were determined of the 22 tetramer STR orthologs isolated from the American wapiti (*Cervus canadensis*) and the mule deer (*Odocoileus hemionus*) available in the literature. Four STR markers were selected by polymorphism screening using the PhastSystem with high resolution PAGE electrophoresis. Of the 10 polymorphic microsatellite loci, two multiple PCR systems each containing 5 STR markers (*DeerPlex I-II*) were constructed. DS-33 (Applied Biosystems) standard matrix fluorophores were used for labelling. Genotyping was performed using the ABI 3130 capillary gel-electrophoresis instrument and the Genemapper® ID 3.2 software (Applied Biosystems).

To estimate the identification power of the markers, the degree of loci polymorphism was characterised by Nei's gene diversity, polymorphism information content (PIC) and the power of discrimination (PD). With the exception of the C229 locus, individual values were high. Random association of alleles within the loci was evaluated by the Hardy-Weinberg equilibrium (HWE) test using the Arlequin v3.11 and Genepop v4.0 softwares. At a significance level of 5%, HWE was rejectable for the T501, C01 and

T108 markers. Independence tests of the loci did not exclude the genetic linkage of C01 and T26 either by statistics (LD test) or by sequence similarity based mapping, therefore the 9 loci not containing the C01 marker were used for forensic calculations.

To test the individual identification potential of *DeerPlex I-II*, the genetic structure of the populations needed to be investigated. The Arlequin software was used to perform a molecular variance analysis (AMOVA), which detected a significant genetic differentiation between the two subpopulations ($F_{ST} = 0,034$); mean inbreeding parameter within the subpopulation was found to be $F_{IS} = 0,037$. The identification power of *DeerPlex I-II* was expressed in average probability of identity (PI_{ave}) and average matching probability (pM) corrected by the F_{IS} and F_{ST} parameters and in the „uniqueness” of the theoretically „most frequent” genetic profile. The probability, that a 9-loci *DeerPlex I-II* profile is unique not higher than $1 - 7,93 \times 10^{-4}$ and the probability that two randomly chosen individuals have identical 9-locus genotypes was found to be $1,0056 \times 10^{-12}$. The results confirmed a very high identification power, similar to that of the Identifiler kit calculated for the Caucasian population of the United States and used for individual identification.

A good example of the forensic application of the marker set is a red deer poaching case. Biological traces collected from the crime scene and the red deer stag trophy confiscated on the suspect's estate were successfully originated from one individual using *DeerPlex I-II* profiles.

The *DeerPlex I-II* systems were used to study the genetic structure of the subpopulations estimating the population origin of individuals. „Assignment test” was used to estimate the reference population of an individual. As 17% of individuals in both populations were classified as not belonging to the reference population, the populations are considered as heterogeneous and hence cannot be used for statistical interpretation of DNA-based evidence in forensic procedures. For this purpose, the total database is recommended to use for forensic calculations. High proportion of individuals (44%) showed overlapping confidence interval of profile frequency values calculated from the two populations. For this reason, determination the population origin of an individual using assignment test would be biased and therefore are not recommended unless the databases are extended in the future.